Can 21-cm observations discriminate between high-mass and low-mass galaxies as reionization sources?

Article (Published Version)

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **423,** 2222–2253 (2012) doi:10.1111/j.1365-2966.2012.21032.x

Can 21-cm observations discriminate between high-mass and low-mass galaxies as reionization sources?

Ilian T. Iliev,[1⋆] Garrelt Mellema,[2] Paul R. Shapiro,[3] Ue-Li Pen,[4] Yi Mao,[3] Jun Koda[3] and Kyungjin Ahn[5]

[1]*Astronomy Centre, Department of Physics and Astronomy, Pevensey II Building, University of Sussex, Falmer, Brighton BN1 9QH*
[2]*Department of Astronomy and Oskar Klein Centre, Stockholm University, Albanova, SE-10691 Stockholm, Sweden*
[3]*Department of Astronomy, University of Texas, Austin, TX 78712-1083, USA*
[4]*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St George Street, Toronto, ON M5S 3H8, Canada*
[5]*Department of Earth Science Education, Chosun University, Gwangju 501-759, Korea*

Accepted 2012 April 2. Received 2012 March 16; in original form 2011 July 25

ABSTRACT

The prospect of detecting the first galaxies by observing their impact on the intergalactic medium (IGM) as they reionized it during the first billion years leads us to ask whether such indirect observations are capable of diagnosing which types of galaxies were most responsible for reionization. We attempt to answer this with new large-scale radiative transfer simulations of reionization including the entire mass range of atomically cooling haloes ($M > 10^8$ M$_\odot$). We divide these haloes into two groups, high-mass, atomically cooling haloes, or HMACHs ($M > 10^9$ M$_\odot$), and low-mass, atomically cooling haloes, or LMACHs ($10^8 < M < 10^9$ M$_\odot$), the latter being susceptible to negative feedback due to Jeans mass filtering in ionized regions, which leads to a process we refer to as self-regulation. We focus here on predictions of the redshifted 21-cm emission, to see if upcoming observations are capable of distinguishing a universe ionized primarily by HMACHs from one in which both HMACHs and LMACHs are responsible, and to see how these results depend upon the uncertain source efficiencies. We find that 21-cm fluctuation power spectra observed by the first-generation Epoch of Reionization 21-cm radio interferometer arrays should be able to distinguish the case of reionization by HMACHs alone from that by both HMACHs and LMACHs, together. Some reionization scenarios, e.g. one with abundant low-efficiency sources versus one with self-regulation, yield very similar power spectra and rms evolution and thus can only be discriminated by their different mean reionization history and 21-cm probability distribution function (PDF) distributions. We also find that the skewness of the 21-cm PDF distribution smoothed with Low Frequency Array (LOFAR)-like resolution shows a clear feature correlated with the rise of the rms due to patchiness. This is independent of the reionization scenario and thus provides a new approach for detecting the rise of large-scale patchiness. The peak epoch of the 21-cm rms fluctuations depends significantly on the beam and bandwidth smoothing size as well as on the reionization scenario and can occur for ionized fractions as low as 30 per cent and as high as 70 per cent. Measurements of the mean photoionization rates are sensitive to the average density of the regions being studied and therefore could be strongly skewed in certain cases. Finally, the simulation volume employed has very modest effects on the results during the early and intermediate stages of reionization, but late-time signatures could be significantly affected.

Key words: radiative transfer – methods: numerical – galaxies: haloes – galaxies: high-redshift – intergalactic medium – dark ages, reionization, first stars.

1 INTRODUCTION

Study of the Epoch of Reionization (EoR) has progressed in recent years in response to a number of new observational developments.

⋆E-mail: I.T.Iliev@sussex.ac.uk

The combination of the cosmic microwave background (CMB) data from the *Wilkinson Microwave Anisotropy Probe* (*WMAP*; Komatsu et al. 2011; Larson et al. 2011) and ever deeper ground-based observations of high-redshift quasi-stellar objects (QSOs), galaxies and gamma-ray bursts (Ouchi et al. 2010; Cucchiara et al. 2011; Kashikawa et al. 2011; Mortlock et al. 2011; Krug et al. 2012) clearly suggests that the reionization process started early and was quite extended in time. However, observations of the effects of the EoR are just beginning to assemble constraints sufficient to diagnose the conditions which brought it about. Ongoing and upcoming observations are expected to put further, much more stringent constraints on the reionization history. The best constraints are likely to result from redshifted 21-cm experiments with the low-frequency radio interferometers Giant Metrewave Radio Telescope (GMRT)[1] (Paciga et al. 2011), Low Frequency Array (LOFAR)[2] (e.g. Harker et al. 2010), Murchison Widefield Array (MWA; Lonsdale et al. 2009)[3] and Precision Array to Probe Epoch of Reionization (PAPER; Parsons et al. 2010). Additional information will come from the Planck satellite (Ade et al. 2011) and measurements of the near-infrared background with Cosmic Infrared Background Experiment (CIBER)[4] and *AKARI*,[5] among others.

The understanding and correct interpretation of these observational results requires detailed modelling. Specific characteristics and features of the observable signatures provide information about different aspects of the EoR. One of the central questions is what are the nature, abundances and physical properties of the ionizing sources. Our purpose in this work is to explore how different types of observations of the EoR might diagnose the nature of the reionization sources.

Our first generation of simulations of inhomogeneous reionization combined cosmological *N*-body simulations of galaxy and large-scale structure formation and the intergalactic density and velocity fields with detailed radiative transfer calculations of the ionizing radiation from every galactic halo whose formation we resolved in a volume large enough to make statistically meaningful predictions of observable consequences of reionization (Iliev et al. 2006a; Mellema et al. 2006b; Iliev et al. 2008a, 2007b; Doré et al. 2007; Iliev et al. 2008b; Harker et al. 2009; Fernandez et al. 2010; Ichikawa et al. 2010). These simulations were in comoving boxes of size $100/h$ (=143 for $h = 0.7$ henceforth) Mpc on a side. Previous simulations had been limited to much smaller volumes, too small to serve this purpose. Earlier simulations of smaller volumes, for example, underestimated the width of the time interval for the global transition of the intergalactic medium (IGM) from neutral to ionized (Gnedin 2000; Ricotti, Gnedin & Shull 2002; Ciardi, Ferrara & White 2003; Sokasian et al. 2003), as well as the amplitude of the kinetic Sunyaev–Zel'dovich (kSZ) fluctuations in the temperature of the CMB from the EoR (Gnedin & Jaffe 2001; Salvaterra et al. 2005). The characteristic size of the intergalactic H II regions during the EoR is expected to reach tens of Mpc (Furlanetto, Zaldarriaga & Hernquist 2004; Furlanetto, McQuinn & Hernquist 2006a; Friedrich et al. 2011) before they grow large enough to overlap. Any fluctuations introduced by this 'patchiness' scale, therefore, require simulation volumes at least this large to model reionization. Moreover, since the first generation of radio observations seeking to detect fluctuations in the brightness temperature of the 21-cm background from the EoR have angular resolution of a few arcminutes, a simulation box size in excess of ∼100 Mpc is necessary to characterize the power spectrum, at wavenumbers as small as $\sim 0.1\ \mathrm{Mpc}^{-1}$ where the initial sensitivity will peak.

Our simulations of a comoving volume $100/h$ (= 143) Mpc on a side were limited, however, by the mass resolution of the *N*-body simulations, to the direct simulation of galactic sources more massive than $\sim 2 \times 10^9\ \mathrm{M}_\odot$. Galactic sources are also possible at lower mass if halo gas can radiatively cool below the halo virial temperature to make star formation possible. This includes haloes above $\sim 10^8\ \mathrm{M}_\odot$, for which collisional excitation of H atoms can radiatively cool the primordial-composition halo gas since the virial temperature is above 10^4 K. We shall refer to these haloes between about 10^8 and $10^9\ \mathrm{M}_\odot$ as low-mass atomic-cooling haloes ('LMACHs'), to distinguish them from the haloes above $10^9\ \mathrm{M}_\odot$, which we shall call high-mass atomic-cooling haloes ('HMACHs'). LMACHs are more numerous than HMACHs at these epochs, so it might be thought that they would dominate reionization. However, unlike the HMACHs, the LMACHs are vulnerable to the negative feedback effects of photoheating if they form inside a pre-existing H II region of the IGM, since the pressure of the IGM would then prevent the intergalactic gas from collapsing gravitationally into their dark matter host haloes (Efstathiou 1992; Shapiro, Giroux & Babul 1994; Quinn, Katz & Efstathiou 1996; Navarro & Steinmetz 1997; Susa & Umemura 2004; Mesinger & Dijkstra 2008; Okamoto, Gao & Theuns 2008). As first emphasized by Shapiro et al. (1994), this limits their contribution to reionization. Haloes of even smaller mass than LMACHs would be even more vulnerable to the negative feedback effects of photoionization heating, since they would, in addition to being prevented from capturing intergalactic gas, photoevaporate whatever interstellar gas they had already accumulated before they were engulfed by reionization (Shapiro, Iliev & Raga 2004; Iliev, Shapiro & Raga 2005). Before that reionization, however, these minihaloes with masses below about $10^8\ \mathrm{M}_\odot$ nevertheless could have formed stars if enough H_2 molecules were present in them. Since their virial temperatures are below 10^4 K, their gas is too cold for collisional excitation and radiative cooling by H atoms, but cooling through the collisionally exciting rotational–vibrational levels of H_2 is possible. However, these molecules were easily dissociated by the rising ultraviolet (UV) background of starlight at energies below the ionization threshold of H atoms, an inevitable by-product of the same stars that contributed to reionization. This tends to limit the contribution of minihaloes to reionization early in the EoR (Haiman, Abel & Rees 2000; Ahn et al. 2009). Previous estimates of the minihalo contribution, as a result, assume that this contribution is smaller than that of the LMACHs and HMACHs, so for now, we shall neglect it, although patchiness in the UV background may make them more important than one might naively think (Ahn et al. 2009).

To investigate the impact of the LMACHs on reionization and its observable properties by direct radiative transfer simulation of reionization, our first generation of simulations boosted the halo mass resolution in order to resolve all the haloes of mass $10^8\ \mathrm{M}_\odot$ and above, but sacrificed volume by simulating in a box of size $37/h$ (= 53 Mpc) on a side (Iliev et al. 2007a). These simulations demonstrated explicitly that reionization in the presence of the LMACHs and their suppression if they formed inside pre-existing H II regions during the EoR was 'self-regulated'. The more LMACHs formed, the more volume and mass of the IGM was ionized, but as this ionized fraction grew, so did the fraction of the total LMACH halo population that formed inside the ionized regions and was suppressed as sources of reionization. This meant that, although

[1] http://gmrt.ncra.tifr.res.in/
[2] http://www.lofar.org/
[3] http://www.mwatelescope.org/
[4] http://physics.ucsd.edu/bkeating/CIBER.html
[5] http://irsa.ipac.caltech.edu/Missions/akari.html

the LMACHs dominated the early phase of reionization, their contribution to reionization eventually saturated, and reionization was finished by the HMACHs, whose abundance rose exponentially over time. An observational consequence of importance that is a possible signature of this process was an early onset but late finish for the EoR, as required to explain the high values of electron-scattering optical depth reported from *WMAP* observations of the large-angle fluctuations in the polarization of the CMB. To make further predictions of observable consequences, however, we need to enlarge the volume of these simulations to at least $100/h = 143$ Mpc on a side, without losing this enhanced mass resolution necessary to resolve the LMACHs. That is the purpose of the new developments we report in this paper.

To accomplish this goal, it was first necessary to improve and advance both our *N*-body and radiative transfer methods in order to be able to simulate halo formation with much higher mass resolution and to transfer the ionizing radiation from a much larger number of sources. Towards this end, both codes had to become massively parallel, running on thousands of computing cores, as well as more efficient. These numerical developments are discussed in more detail in Section 2 below and in Iliev et al. (2008c).

While the work described here follows naturally from our own previous work as described above, it also differs substantially from other work in the literature to date involving large-scale radiative transfer simulations of reionization. A full account of that literature is beyond the scope of this paper, but we will mention a few points to distinguish the current work. Our *N*-body simulations resolve all galactic halo sources of 10^8 M$_\odot$ and above, in a comoving box as large as $114/h = 163$ Mpc on a side. We post-process the density field of the IGM and the galactic halo source populations derived from these *N*-body simulations by performing a detailed, ray-tracing calculation on a grid of 256^3 cells. The simulations described in McQuinn et al. (2007) were based upon post-processing *N*-body simulations with halo mass resolution above 10^9 M$_\odot$, in a box $65.6/h = 94$ Mpc on a side, a volume which is five times smaller than ours, also on a grid of 256^3 cells. McQuinn et al. (2007) were, thus, unable to treat explicitly the halo mass range below 10^9 M$_\odot$, which is subject to suppression by the negative feedback effects of reionization, but they did include a semi-analytical 'subgrid' approximation for the contribution from the unresolved, smaller mass haloes, including some feedback effects. Trac & Cen (2007) considered a $50/h = 71$ Mpc box, hence, an order of magnitude smaller volume than ours, and 180^3 cells for radiative transfer, but their halo mass resolution was similar to ours. They did not consider the feedback effects of reionization on the small-mass galactic halo sources as we do here. Shin, Trac & Cen (2008) subsequently applied this method to simulate a volume of $100/h = 143$ Mpc on a side (2/3 of our volume), with similar halo mass resolution, on a radiative transfer grid of 360^3 cells, to study the structure of the patchy ionization field during the EoR, again without considering feedback and doing only a single simulation, without considering any variation of the (highly uncertain) reionization parameters. Recently, the codes of McQuinn et al. (2007) and Trac & Cen (2007) were compared with each other in Zahn et al. (2011), by examining results for the ionization fields and 21-cm brightness temperature fluctuation statistics, for reionization simulations advanced part-way through the epoch of reionization, to the 72 per cent ionized point, based upon post-processing a previously simulated input density field in a box of $100/h = 143$ Mpc on a side, smoothed to a radiative transfer grid with 256^3 cells, with halo mass resolution of 10^8 solar masses, without any feedback or back-reaction on the halo sources. They also considered a single reionization scenario, with no parameter variation. Finally, Aubert & Teyssier (2010) simulated radiative transfer in several different boxes, as large as $100/h = 143$ Mpc, by post-processing a density field and galaxy population derived from separate simulations that combined *N*-body dynamics and hydrodynamics on a 1024^3 grid, but with minimum resolved galaxy masses in that case as large as 8×10^9 M$_\odot$. No feedback from reionization on galactic sources was considered. While there are other distinctions of detail between these other simulations and ours, we have listed these above to make it clear that our current paper will describe simulations and results which are new, *both* because they are based on a different methodology, applied in greater depth than previously to predict observables like the 21-cm background from the EoR, *and* because they are on an unprecedented scale.

The rest of this paper is organized as follows. In Section 2 we present our codes, numerical methods and simulations. In Section 3 we discuss our results on the formation of early cosmic structures. In Section 4 we present our results on the basic reionization features, reionization history, integrated electron-scattering optical depth and geometry of ionized patches. The observational signatures derived from our simulations are presented and discussed in Section 5. Our conclusions are summarized in Section 6. Finally, in Appendix A we present the (physically less realistic) cases of reionization by rare, massive sources, while in Appendix B we discuss the more technical point of the dependence of our results on the Jeans suppression threshold for low-mass sources.

2 SIMULATIONS

Our basic methodology has been previously described in Iliev et al. (2006a, 2007a) and Mellema et al. (2006b). Due to the much larger scale of our current simulations compared to our previous ones, both our structure formation and our radiative transfer code had to be significantly developed and re-designed, in particular to allow their massive parallelization on distributed-memory machines. In this section we present our new set of simulations, along with a summary of our methods and parallel code scaling to large number of computing cores.

2.1 *N*-body simulations

We start by performing very high-resolution *N*-body simulations of the formation of high-redshift structures. We use the CUBEP^{3}M *N*-body code[6] which evolved from the particle-mesh (PM) code PMFAST (Merz, Pen & Trac 2005). In CUBEP^{3}M several important new features were introduced in comparison with these previous codes. The first one is the addition of a short-range direct particle–particle force, making it a P^{3}M (particle–particle–particle-mesh) code. This significantly improves its spatial resolution and accuracy at small scales compared to PM codes. A second important new development is that CUBEP^{3}M is now a massively parallel code which can run efficiently on either distributed- or shared-memory machines. This is achieved through cubical equal-volume domain decomposition and a hybrid Message Passing Interface (MPI) and OpenMP approach (see Iliev et al. 2008c for more details). CUBEP^{3}M scales well (with 'weak' scaling, whereby the execution time rises proportionally to the problem size) up to thousands of processors, as shown in Fig. 1 (left) and has to date been run on up to 21 976 computing cores, following up to 5488^3 particles (Iliev et al. 2010). These scaling

[6] http://www.cita.utoronto.ca/mediawiki/index.php/CubePM, for description of the code; see also Iliev et al. (2008c).

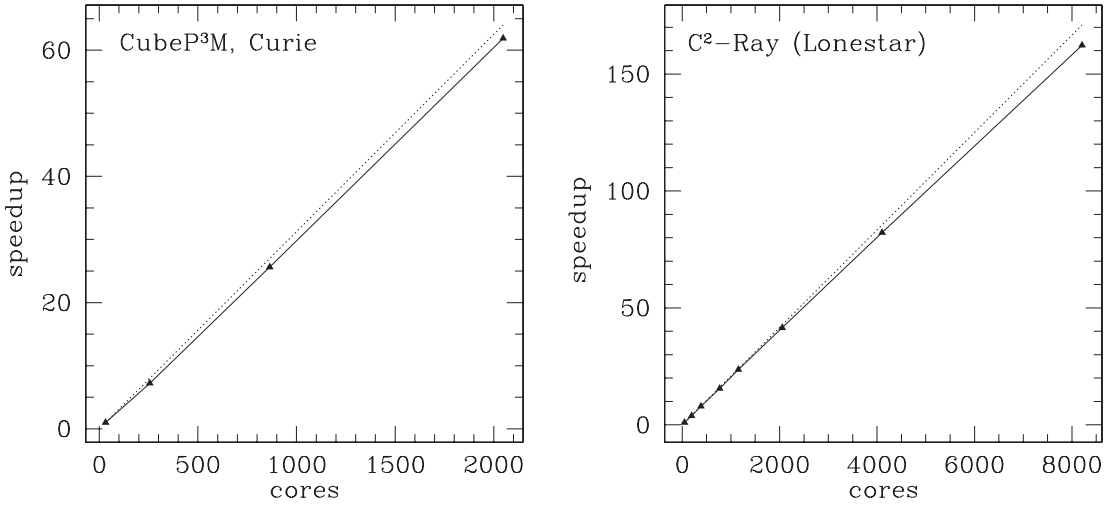

Figure 1. Left: scaling of the CUBEP^{3}M code. Plotted are the code speed-up versus the number of computational cores used. Both quantities are normalized to the smallest run in each case. Right: scaling of the C^2-RAY code. Plotted are the code speed-up versus the number of computational cores used, again normalized to the smallest of the three runs compared. The dashed line indicates the ideal weak scaling for each case.

Table 1. *N*-body simulation parameters. Background cosmology is based on the *WMAP* 5-year results.

Box size	$N_{\rm part}$	Mesh	Spatial resolution	$m_{\rm particle}$	$M_{\rm halo,min}$
37 h^{-1} Mpc = 53 Mpc	1024^3	2048^3	1.81 kpc h^{-1}	5.05×10^6 M$_\odot$	1.01×10^8 M$_\odot$
64 h^{-1} Mpc = 91 Mpc	1728^3	3456^3	1.85 kpc h^{-1}	5.44×10^6 M$_\odot$	1.09×10^8 M$_\odot$
74 h^{-1} Mpc = 106 Mpc	2048^3	4096^3	1.81 kpc h^{-1}	5.05×10^6 M$_\odot$	1.01×10^8 M$_\odot$
114 h^{-1} Mpc = 163 Mpc	3072^3	6144^3	1.86 kpc h^{-1}	5.47×10^6 M$_\odot$	1.09×10^8 M$_\odot$

tests were run on the Texas Advanced Computing Center (TACC) computer, Lonestar, and on the currently available portion of the European Petascale computer under development in France, Curie, at CEA, part of the Partnership for Advanced Computing in Europe (PRACE). Our results show almost perfect scaling of CUBEP^{3}M, within 3 per cent of the ideal one (dashed line), for up to 2048 cores.

We performed a series of *N*-body cosmic structure formation simulations (summarized in Table 1) with an increasing simulation box size, from 37 h^{-1} Mpc (53 Mpc) up to 114 h^{-1} Mpc (163 Mpc), but all with a fixed spatial and mass resolution. These *N*-body simulations were run on a range of core numbers from 256 (1024^3 particles) up to 2048 cores (3072^3 particles). The force smoothing length is fixed to 1/20 of the mean interparticle spacing, or 1.8 h^{-1} kpc. The largest of these simulations follows a volume which is 50 per cent larger than the largest structure formation simulation performed previously at similar resolution.

The *N*-body simulations required between 4100 (for 1024^3 particles) and 159 000 (for 3072^3 particles) computing hours (computing cores $\times$ wall-clock hours) on the TACC computer Ranger (SunBlade x6420 with AMD $\times$86-64 Opteron Quad Core, 2.3 GHz, 9.2 GFlops/core Barcelona processors and Infiniband networking). The particle mass is 5×10^6 M$_\odot$, which guarantees that all atomically cooling haloes ($M > 10^8$ M$_\odot$) are resolved with at least 20 particles. We use a spherical overdensity halo finder with an overdensity parameter fixed to 178 of mean density (note that $\bar{\rho}_M \neq \rho_{\rm crit}$ except for very high redshift). We note that compared to the frequently used friends-of-friends (FOF) halo finder, the spherical overdensity halo finders typically yield slightly different halo masses (White 2001; Reed et al. 2007; Tinker et al. 2008). However, these differences (typically of the order of 10–20 per cent) are relatively small compared to the much larger uncertainties due to the still weakly constrained photon production efficiencies (see Section 2.2 below). Our series of *N*-body simulations with an increasing size allowed us to test the convergence of our results with computational box size.

The background cosmology is based on the *WMAP* 5-year data combined with constraints from baryonic acoustic oscillations and high-redshift supernovae ($\Omega_{\rm M} = 0.27$, $\Omega_\Lambda = 0.73$, $h = 0.7$, $\Omega_{\rm b} = 0.044$, $\sigma_8 = 0.8$and $n = 0.96$). The linear power spectrum of density fluctuations was calculated with the code CAMB (Lewis, Challinor & Lasenby 2000). Initial conditions were generated using the Zel'dovich approximation at sufficiently high redshift ($z_i = 300$) to ensure against numerical artefacts (Crocce, Pueblas & Scoccimarro 2006).

2.2 Radiative transfer simulations

The radiative transfer simulations are performed with our code C^2-RAY (Conservative Causal Ray-Tracing) (Mellema et al. 2006a). The method is explicitly photon-conserving in both space and time for individual sources and approximately (to a good approximation) photon-conserving for multiple sources, which ensures correct tracking of ionization fronts without loss of accuracy, independent of the spatial and time resolution, with corresponding great gains in efficiency. The code has been tested in detail against a number of exact analytical solutions (Mellema et al. 2006a), as well as in direct comparison with a number of other independent radiative transfer methods on a standardized set of benchmark problems (Iliev et al. 2006b, 2009). The ionizing radiation is ray-traced from every source to every grid cell using the short characteristics method, whereby the neutral column density between the source and a given cell is given by interpolation of the column densities of the previous cells which lie closer to the source, in addition to the neutral column

density through the cell itself. The contribution of each source to the local photoionization rate of a given cell is first calculated independently, after which all contributions are added together and a non-equilibrium chemistry solver is used to calculate the resulting ionization state. Ordinarily, multiple sources contribute to the local photoionization rate of each cell. Changes in the rate modify the neutral fraction and thus the neutral column density, which in turn changes the photoionization rates themselves (since either more or less radiation reaches the cell). An iteration procedure is thus called for in order to converge to the correct, self-consistent solution. While our basic methodology remains essentially as described in Mellema et al. (2006a), our C^2-RAY code has been thoroughly re-written in FORTRAN90, made more flexible and modular and parallelized for distributed-memory machines. In terms of a parallelization strategy, due to the causal nature of the ray-tracing procedure (i.e. the state of each cell can be calculated only after all previous cells closer to the source are done) it is not possible to employ domain decomposition (except for a limited one, into octants, see below), although other approaches exist which seek ways to overcome this limitation (Nakamoto, Umemura & Susa 2001; Rijkhorst et al. 2006). Instead, the main code loop over the sources of ionizing radiation is done in a massively parallel fashion. Each MPI node has a copy of the density field and receives a number of sources whose radiation is to be traced through the grid. For the large-scale cosmological reionization problem there are typically hundreds of thousands to millions of sources; thus, our code scales well up to tens of thousands of cores at least (see the next section). For problems with (relatively) low number of ionizing sources such a parallelization strategy would be inefficient, but such problems are not sufficiently computationally intensive to require such massive parallelization and could, instead, be solved on a smaller number of nodes, or even in serial. A similar situation occurs for the initial steps of the simulations presented below, when the cosmological structure formation is not yet much advanced; thus, only a few to few tens of haloes form. However, their number increases exponentially over time, quickly reaching thousands, and then tens and hundreds of thousands. We therefore start our simulations on a small number of cores (typically 32), raising this to thousands of cores as more sources form.

As mentioned above, a limited domain decomposition on to octants is possible for our method, since those are independent of each other within the short-characteristic ray-tracing framework. We use this to (optionally) improve the memory efficiency of the code by doing the grid octants in parallel within each MPI node using OpenMP multi-threading. This way each MPI node needs only one copy of the grid, which is shared amongst the cores within the node.

The radiative transfer problem size scales proportionally to both the grid size and the number of sources. Results shown in Fig. 1 (right) demonstrate almost perfect scaling, within ∼10 per cent from the ideal one, for up to 8192 cores.

The *N*-body simulations discussed above provide us with the spatial distribution of cosmological structures and their evolution in time. We then use this information as input to a full 3D radiative transfer simulations of the reionization history, as follows. We saved series of time slices, both particle lists and halo catalogues from redshift 50 down to 6, uniformly spaced in time, every $\Delta t = 11.53$ Myr, a total of 76 slices. Based on the particle distribution at each redshift, we used smoothed particle hydrodynamics (SPH)-style smoothing scheme using the nearest neighbours (to be described in detail in a companion paper Koda et al., in preparation) to produce regular-grid density and bulk velocity fields at the radiative transfer resolution of 256^3 cells.

All identified haloes are potential sources of ionizing radiation, with a photon production rate per Myr, $\dot{N}_\gamma$, proportional to their mass, M:

$$\dot{N}_\gamma = \frac{f_\gamma M \Omega_\mathrm{b}}{\Delta t\, \Omega_0 m_\mathrm{p}}, \quad (1)$$

where m_p is the proton mass and $f_\gamma = f_\mathrm{esc} f_\star N_\star$ is an ionizing photon production efficiency parameter which includes the efficiency of converting gas into stars, f_*, the ionizing photon escape fraction from the halo into the IGM, f_esc, and the number of ionizing photons produced per stellar atom, $N_\star$. The latter parameter depends on the assumed initial mass function (IMF) for the stellar population and varies between 4000 and ∼100 000. Haloes were assigned different luminosities according to whether their mass was above (HMACHs) or below (LMACHs) $10^9\,\mathrm{M}_\odot$ (but above $10^8\,\mathrm{M}_\odot$, the minimum resolved halo mass). LMACHs are assumed to be suppressed within ionized regions (for ionization fraction higher than 10 per cent), through Jeans mass filtering, as discussed in Iliev et al. (2007a).

We note that while previously we used the factor f_γ to characterize the source efficiencies, here we define a slightly different factor, g_γ, that is given by

$$g_\gamma = f_\gamma \left(\frac{10\ \mathrm{Myr}}{\Delta t}\right), \quad (2)$$

where Δt is the time between two snapshots from the *N*-body simulation. The new factor g_γ has the advantage that it is independent of the length of the time interval between the density slices, and as such it allows a direct comparison between runs with different Δt. For readers' convenience we listed the values of both parameters in Table 2. We also note that the specific numerical values of the efficiency parameters are strongly dependent on the background cosmology adopted and the minimum source halo mass. Therefore, parameter values for simulations based on different underlying cosmology and resolution should not be compared directly, but would require a cosmology and resolution-dependent conversion coefficients to achieve the same reionization history.

Our full simulation notation reads $Lbox\_gI\_J(S)(K)$ (the bracketed quantities are listed only when needed), where '*Lbox*' is the simulation box size in Mpc, '*I*' and '*J*' are the values of the g_γ factor for HMACHs and LMACHs, respectively, the symbol '*S*' means that the small sources are suppressed within already-ionized regions and '*K*' indicated the ionized fraction threshold for a given radiative transfer cell above which this suppression occurs for haloes residing in that cell, which is 0.1 if not listed explicitly and raised to 0.9 or 0.5 for cases $K = 9$ and $K = 5$, respectively (see below for details). For example, 53Mpc_g8.7_130S indicates that large sources have an efficiency of $g_\gamma = 8.7$, while small sources have an efficiency of $g_\gamma = 130$ and are suppressed in ionized regions.

We have performed series of radiative transfer simulations with varying underlying assumptions about the source efficiencies and the suppression conditions imposed on the LMACHs, as summarized in Table 2. For our radiative transfer simulations we use the data from the largest *N*-body box, $114/h = 163$ Mpc, and the smallest one, $37/h = 53$ Mpc. The former volume is sufficiently large to faithfully represent the reionization observables, while the latter one affords much faster and computationally cheaper simulations, which allows us to explore a wider parameter space. These two very different computational volumes also allow us to investigate resolution effects and evaluate which features of reionization and observable signatures are sensitive to the box size and which are less so. We label all runs by a short label (listed in the first column of Table 2) for more compact notation. Large-box runs are labelled L1–L3, while small-box runs are labelled S1–S9.

Table 2. Reionization simulation parameters and global reionization history results. All runs use background cosmology based on the *WMAP* 5-year results. Cases S6–S9 are used to test certain aspects of our model, see Appendices A and B.

Label	Run	Box size (cMpc)	$g_\gamma(f_\gamma)^a$ HMACH	$g_\gamma(f_\gamma)$ LMACH	Mesh	Supp.	Min. source ($M_\odot$)	Min. unsupp. halo ($M_\odot$)	τ_{es}	$z_{10\,per\,cent}$	$z_{50\,per\,cent}$	$z_{90\,per\,cent}$	z_{ov}
L1	163Mpc_g8.7_130S	163	8.7 (10)	130 (150)	256^3	Yes	10^8	10^9	0.080	13.3	9.4	8.6	8.3
L2	163Mpc_g1.7_8.7S	163	1.7 (2)	8.7 (10)	256^3	Yes	10^8	10^9	0.058	9.9	7.6	6.9	6.7
L3	163Mpc_g21.7_0	163	21.7 (25)	0 (0)	256^3	No	2.2×10^9	2.2×10^9	0.070	10.3	9.1	8.6	8.4
S1	53Mpc_g8.7_130S	53	8.7 (10)	130 (150)	256^3	Yes	10^8	10^9	0.084	13.6	9.8	8.9	8.5
S2	53Mpc_g1.7_8.7S	53	1.7 (2)	8.7 (10)	256^3	Yes	10^8	10^9	0.059	10.0	7.7	6.9	6.7
S3	53Mpc_g8.7_130	53	8.7 (10)	130 (150)	256^3	No	10^8	10^8	0.131	15.6	13.9	13.2	12.9
S4	53Mpc_g0.4_5.3	53	0.35 (0.4)	5.3 (6)	256^3	No	10^8	10^8	0.078	11.7	9.7	8.9	8.6
S5	53Mpc_g10.4_0	53	10.4 (12)	0	256^3	No	10^9	10^9	0.071	10.5	9.1	8.5	8.3
S6	53Mpc_g8.7_130S9	53	8.7 (10)	130 (150)	256^3	Yes[c]	10^8	10^9	0.111	14.9	12.6	10.7	9.5
S7	53Mpc_g8.7_130S5	53	8.7 (10)	130 (150)	256^3	Yes[c]	10^8	10^9	0.089	13.9	10.1	9.0	8.6
S8	53Mpc_uvS_1e9	53	Variable[b]	0	256^3	No	10^9	10^9	0.084	13.7	9.7	8.9	8.5
S9	53Mpc_uvS_1e10	53	Variable[b]	0	256^3	No	10^{10}	10^{10}	0.080	12.2	9.8	8.9	8.5

[a] f_γ is related to g_γ by equation (2) with $\Delta t = 11.53$ Myr. [b]See Fig. A1 and discussion in Appendix A. [c]Employing a different suppression criterion, see Appendix B.

Our fiducial runs, to which all the others will be compared, are 163Mpc_g8.7_130S (L1) and the companion small-box one with same source efficiencies, 53Mpc_g8.7_130S (S1). These parameters yield a relatively early overlap (which we define as the time when the mass-weighted ionized fraction reaches 99 per cent) and high electron scattering optical depth. The second set of simulations, 163Mpc_g1.7_8.7S (L2) and 53Mpc_g1.7_8.7S (S2), are in the opposite limit, which assumes considerably lower efficiencies for both types of sources and as a consequence serves as a model for a late-overlap, extended, more photon-poor reionization scenario. These two cases are designed to roughly bracket the range of observationally allowed reionization scenarios.

Our third large-box simulation, 163Mpc_g21.7_0 (L3) is equivalent to our previous large-box simulations without self-regulation presented in Iliev et al. (2008a), except for the updated background cosmology and the source photon production efficiencies, adjusted here to yield the same overlap epoch as our fiducial case L1. Therefore, L1 and L3 share the underlying density structures and sources (apart from the different minimum mass cut-offs) and hence a head-to-head comparison yields the effects of the presence of LMACHs and Jeans mass filtering.

The rest of our cases, S3 to S9, test various aspects of the reionization source modelling. Simulation S3 is an extreme case which has the same source efficiencies as our fiducial case L1, but assumes no-suppression occurs. Naturally, this results in a very early reionization and very high integrated optical depth, $\tau_{es} = 0.131$, which is well outside the *WMAP*5 1σ range of $\tau_{es} = 0.084 \pm 0.016$. Simulation 53Mpc_g0.4_5.3 (S4) is again without suppression, but here we tuned down the efficiencies of both types of sources so as to achieve approximately the same overlap epoch as in our fiducial case L1. In simulation 53Mpc_g10.4_0 (S5) we assume that there are no LMACHs ($M < 10^9\,M_\odot$) at all and we adjust the photon efficiency of the remaining sources to again reach overlap at roughly the same epoch as in the fiducial case.

Additionally, we consider two scenarios which have exactly the same time-dependent ionizing photon emissivity (and therefore almost identical reionization history) as our fiducial case S1, but with a higher minimum source mass of $10^9\,M_\odot$ (53Mpc_uvS_1e9; S8) and $10^{10}\,M_\odot$ (53Mpc_uvS_1e10; S9). The fixed ionizing photon emissivity results in unphysically high early source luminosities and we consider them primarily in order to illustrate the effect of redistributing the full luminosity of all sources over the massive ones only, similar to the models adopted in some recent work (Baek et al. 2009; Thomas et al. 2009). These simulations and some illustrative results from them are discussed in Appendix A.

Finally, in order to evaluate the robustness of our source suppression model, we consider two more scenarios, 53Mpc_g8.7_130S9 (S6) and 53Mpc_g8.7_130S5 (S7), whereby we raise the ionization threshold for LMACH suppression to $x_{threshold} = 0.9$ and 0.5, respectively, from our fiducial threshold of $x_{threshold} = 0.1$. Since this is a more technical study we present its results separately, in Appendix B.

The radiative transfer simulations presented in this work typically required ∼0.5–1 million computing hours (163 Mpc boxes) and ∼10–30 thousand computing hours (53 Mpc boxes), depending on the specific set of source parameters we adopted.

3 RESULTS: EARLY STRUCTURE FORMATION

In Fig. 2 (left: full box, right: zoomed subvolume) we show a slice of the density field and haloes at redshift $z = 6$ from our $114\,h^{-1}$ Mpc (163 Mpc), 3072^3-particle N-body simulation. The structure formation is already well-advanced and strongly non-linear at sub-Mpc scales. The very first resolved ($M_{halo} > 10^8\,M_\odot$) halo in this volume forms at $z = 31$, while the first insuppressible halo (HMACH; $M_{halo} > 10^9\,M_\odot$) forms at $z = 21$. By $z = 6$ there are over 20.5 million collapsed haloes (corresponding to a mean number density of $n = 13.8\,h^3\,\mathrm{Mpc}^{-3}$), of which ∼18.7 million LMACHs ($M_{halo} < 10^9\,M_\odot$; $n = 12.6\,h^3\,\mathrm{Mpc}^{-3}$) and ∼2 million HMACHs ($M_{halo} > 10^9\,M_\odot$; $n = 1.2\,h^3\,\mathrm{Mpc}^{-3}$). The haloes are strongly clustered at all times, more so going to higher redshifts, when they are ever rarer. The halo abundances are usually described in terms of $\nu = \delta_c D_+/\sigma(0, M)$, where δ_c is the linear density contrast corresponding to the moment of collapse of a top-hat density perturbation, $\sigma(0, M)$ is the present variance of the density fluctuations corresponding to the mass scale M, and D_+ is the growth factor of the density fluctuations. The halo masses corresponding to $\nu = 1$ (most common, M_*, haloes), $\nu = 2$, 3 (rare haloes) and $\nu = 5$ (extremely rare haloes) are shown in Fig. 3 (left). Clearly, before $z \sim 6$, within the redshift range of interest here, there are no 1σ haloes and thus all haloes are statistically rare. More specifically, atomically cooling haloes are 2σ–3σ at the low end of the

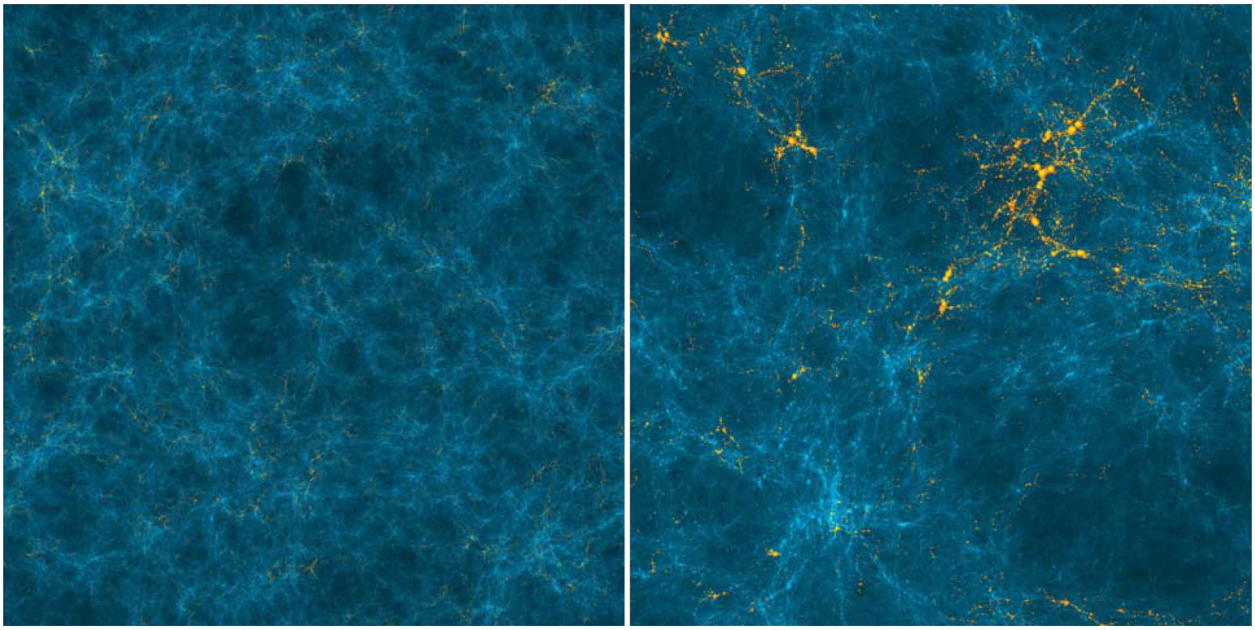

Figure 2. Left: slice of the Cosmic Web at redshift $z = 6$ from our CUBEP^{3}M simulation with 3072^3 particles (29 billion) on a 6144^3 fine grid in a comoving volume of $114/h = 163$ Mpc on a side. Shown are the dark matter density (blue) and haloes (in actual size; yellow). Image resolution is 6144×6144, the slice is $1/h = 1.42$ Mpc thick. Right: zoomed-up region (25.76×25.76 Mpc) of the same image.

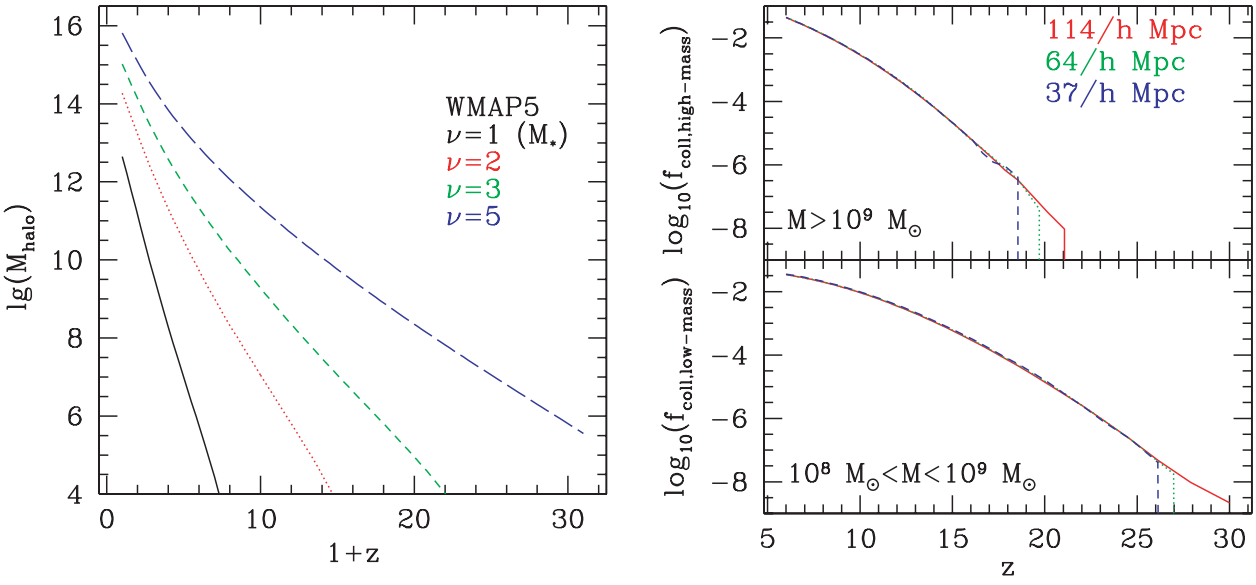

Figure 3. Left: halo mass versus redshift for $\nu = \delta_c D_+/\sigma(0, M) = 1$ (i.e. M_*, typical halo; black, solid), $\nu = 2$ (red, dotted), $\nu = 3$ (green, short-dashed) and $\nu = 5$ (blue, long-dashed). Right: redshift evolution of the collapsed fraction (fraction of the total mass found inside collapsed haloes) of HMACHs (top) and LMACHs (bottom).

redshift interval and as rare as 5σ at early times. The evolution of the collapsed fractions in HMACHs and LMACHs is shown in Fig. 3 (right). The collapsed fractions start very low and rise exponentially at early times when the haloes are very rare. The collapsed fraction in LMACHs reaches 10^{-3} and 1 per cent at $z = 14$ and $z = 9.8$, respectively. After that point it starts to level off as LMACHs start to become less rare (2σ or less) and their collapsed fraction reaches 3.4 per cent by $z = 6$. The collapsed fraction in HMACHs rises steeply all the way to $z = 6$, eventually reaching 4.25 per cent. There is only a modest departure from the exponential growth, reflecting the fact that they remain quite rare throughout this period. The simulation volume has essentially no effect on the derived collapsed fractions, indicating numerical convergence on that quantity at fixed mass resolution. The only exception to this is that at very early times ($z > 26$ for the LMACHs and $z > 17$ for the HMACHs) the corresponding halo populations are so rare ($\sim 5\sigma$ in each case) that Poisson noise (i.e. cosmic variance, due to the smaller volume) affects the results. For example, the first resolved haloes form at

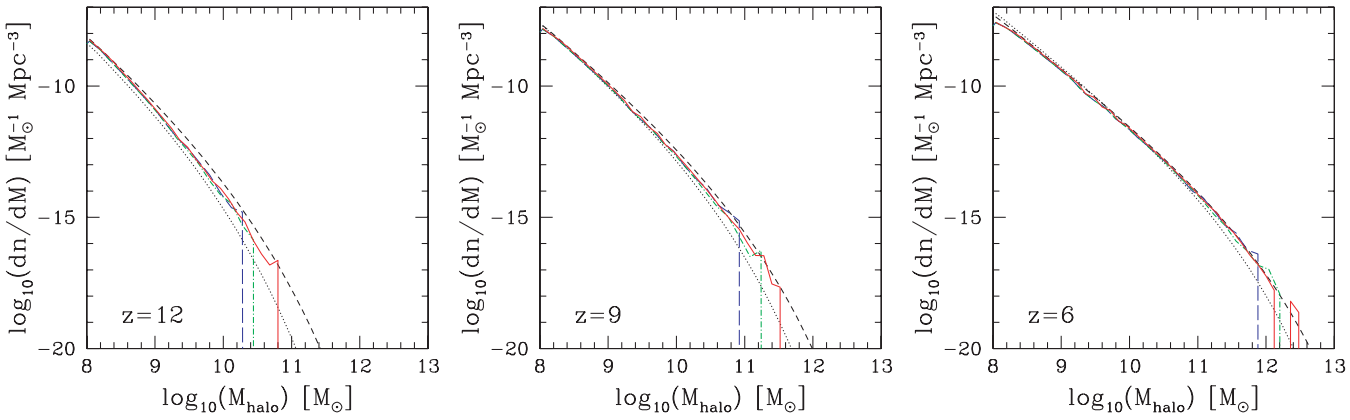

Figure 4. Simulated halo mass function at high-z derived from $114/h = 163$ Mpc box (solid, red), $74/h = 106$ Mpc box (dot–dashed, green), $37/h = 53$ Mpc box (long-dashed, blue) at (left to right) $z = 12$, 9 and 6. Also shown are the Press–Schechter (dotted, black) and Sheth–Tormen (short-dashed, black) analytical mass functions.

$z = 31$ in the $114\,h^{-1} = 163$ Mpc box, but only at $z = 26$ in the $37\,h^{-1} = 53$ Mpc box. As soon as there is sufficient statistics for any given volume the collapsed fractions converge.

The halo mass functions derived from our simulations are shown in Fig. 4 for a range of redshifts, $z = 12$–6, along with the Press–Schechter (PS; Press & Schechter 1974) and Sheth–Tormen (ST; Sheth & Tormen 2002) analytical mass functions. The halo abundances at all redshifts fall between those two analytical predictions. PS always underpredicts the abundances of massive, rare haloes, while ST overpredicts them (see also Cohn & White 2008; Dalal et al. 2008). With time ST becomes a better match to the numerical results. This broadly agrees with previous results on the high-redshift mass functions (Iliev et al. 2006a; Lukić et al. 2007; Reed et al. 2007; Cohn & White 2008).

In Fig. 5 we plot the total matter density field power spectra $\Delta_\rho = (k^3 P(k)/2\pi^2)^{1/2}$ at two representative redshifts for our largest (114 Mpc h^{-1}) and smallest (37 Mpc h^{-1}) boxes. Power spectra were calculated by interpolating the N-body particles using a cloud-in-cell scheme on to the fine grid of the CUBEP^{3}M code, with 6144^3 and 2048^3 cells, respectively. There is a close agreement between the two cases, apart from the expected variance at scales close to the box size. This shows that there is no missing density fluctuation power in the small box, except for the scales at or above the box size.

Finally, in Fig. 6 we show the halo–halo bias, calculated as the ratio of the halo autocorrelation power spectrum divided by the density field one, i.e. $b_{\rm hh} = \Delta_{\rm hh}/\Delta_\rho$. This measures the clustering of the dark matter haloes with respect to the underlying matter density field at redshifts $z = 12$ and 9, roughly corresponding to the early and advanced stages of reionization, respectively. Because of the relative rarity of all haloes studied here, their clustering is quite strong, particularly at the smallest resolved scales ($k \sim 100\,h\,{\rm Mpc}^{-1}$), where it is of the order of 100 for the lowest mass haloes and is as high as $\sim 10^4$ for the most massive haloes. At these small scales the bias is strongly non-linear. At large scales the bias factors asymptote to a (mass-dependent) constant – the large-scale linear bias (Mo & White 1996). The transition between the large-scale linear bias and the non-linear one occurs around $k \sim 1$

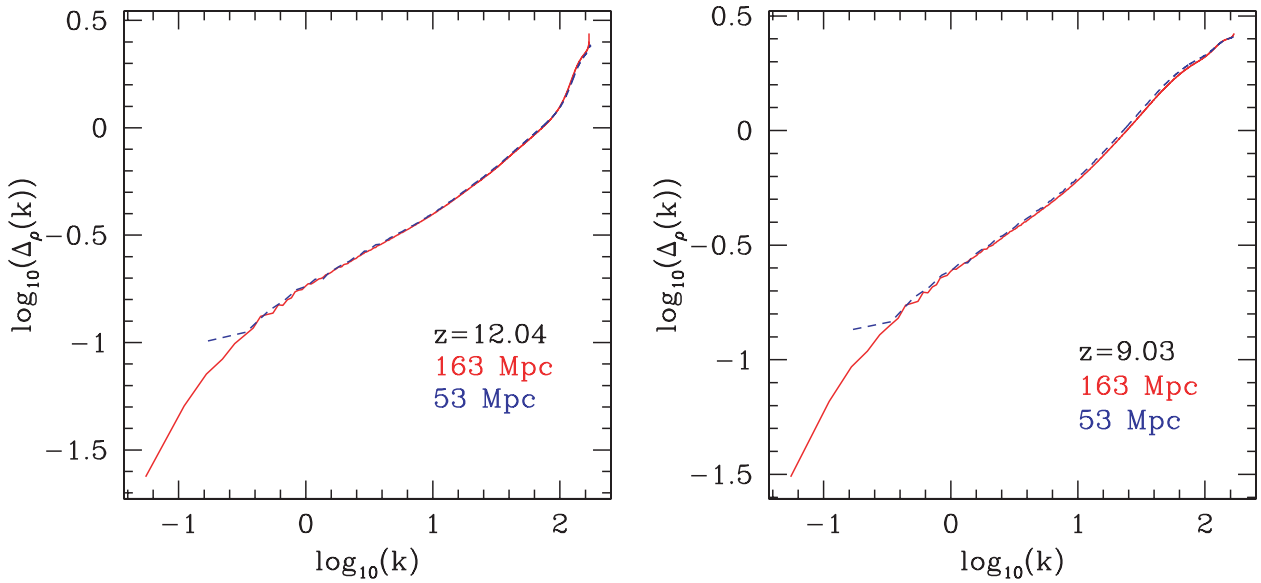

Figure 5. The power spectra of the density fields, $\Delta_\rho = (k^3 P(k)/2\pi^2)^{1/2}$, for the $114/h = 163$ Mpc box run (red, solid) and the $37/h = 53$ Mpc box run (blue, dashed) at redshifts $z = 12$ (left) and 9 (right).

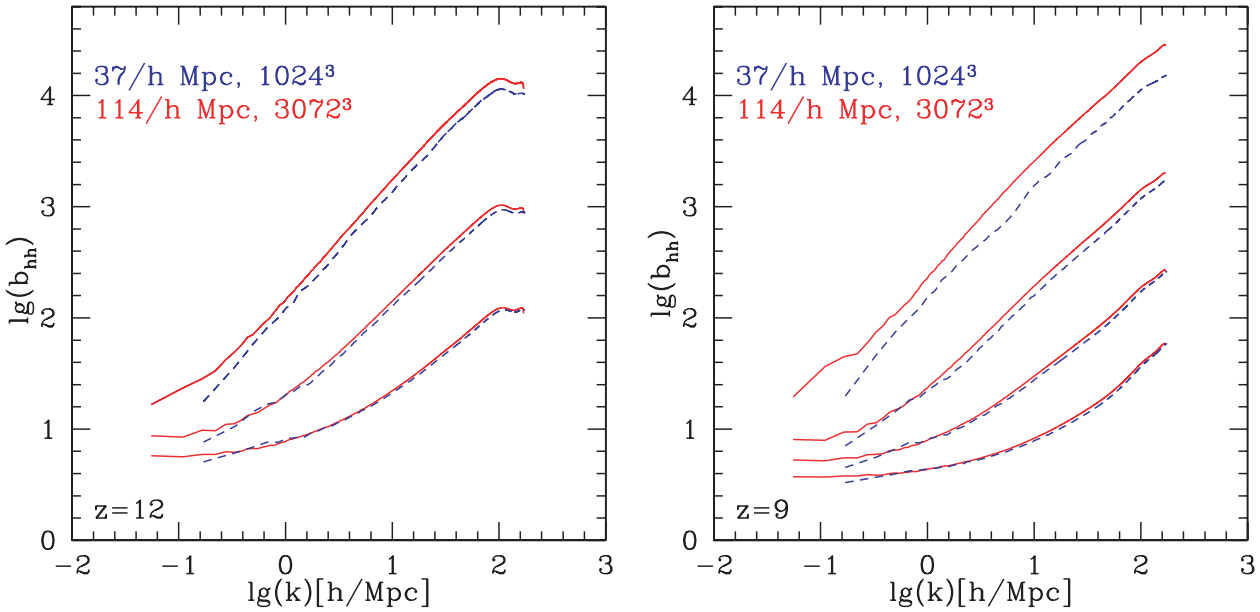

Figure 6. The halo bias, $b_{\rm hh} = \Delta_{\rm hh}/\Delta_\rho$, for the $114/h = 163$ Mpc (red, solid) and $37/h = 53$ Mpc box (blue, dashed) runs at redshifts $z = 12.04$ (left) and 9.03 (right). Lines are for haloes binned by decades of mass (bottom to top curve) $10^8 \leq M_{\rm halo} < 10^9\,{\rm M}_\odot$ (LMACHs), $10^9 \leq M_{\rm halo} < 10^{10}\,{\rm M}_\odot$, $10^{10} \leq M_{\rm halo} < 10^{11}\,{\rm M}_\odot$ and $10^{11} \leq M_{\rm halo} < 10^{12}\,{\rm M}_\odot$ (HMACHs all).

for the lowest mass haloes at $z = 9$, rising to $k \sim 0.1$–0.5 for the larger haloes and/or higher redshifts. The largest, rarest haloes ($M > 10^{10}\,{\rm M}_\odot$ at $z = 12$, $M > 10^{11}\,{\rm M}_\odot$ at $z = 9$) show a roughly linear log–log $b(k)$ relation and never asymptote to the linear bias value within the k-range covered in our simulation. Comparing the results from the two simulation volumes, $114h^{-1}$ Mpc versus $37h^{-1}$ Mpc, we see that the bias for the LMACHs and the lower mass bins for the HMACHs agree very well, indicating good numerical convergence with the simulation volume. The bias for the largest mass haloes is underestimated in the small volume due to the small number statistics and the suppressed long-wavelength density fluctuations in that case.

4 RESULTS: BASIC FEATURES OF THE SELF-REGULATED REIONIZATION

4.1 The effects of self-regulation

We start by comparing our fiducial simulation L1 against the simulation L3, where the latter simulation is equivalent to our previous large-box simulations which were without LMACHs and self-regulation (Iliev et al. 2008a). The resulting reionization histories and integrated electron-scattering optical depths are shown in Fig. 7. As expected, the self-regulated model yields a much more gradual and extended reionization history, which starts with the formation of the first atomically cooling halo sources at $z \sim 31$ versus the much later start, at $z \sim 20$ in L3, due to the much larger, rarer sources in the latter case, which accordingly form later. The exponential rise in the numbers of the high-mass sources yields a steep, power-law-like reionization history [reasonably well fitted by $\log(x_{\rm m}) \approx -1.226z + 10.41$ for $x_{\rm m} > 0.03$, rising somewhat steeper than this earlier on] when it is driven solely by those sources, while in the self-regulated case the steep initial rise becomes much more gradual when the self-regulation first kicks in, around $z \sim 16$, when the high-efficiency LMACHs are massively suppressed and thereby gradually give way to the less efficient HMACHs which come to dominate at the later stages of reionization. However, we note that even with self-regulation the reionization history remains monotonic and no plateaus, let alone double reionization, ever occur. The mass weighted over volume-ionized fraction (upper panel) is always lower in the self-regulated case, indicating that reionization has a less pronounced inside-out character, i.e. ionized regions are less correlated with the highest density peaks in this case since reionization is driven by a wider range of sources, including low-mass, less biased ones. The integrated electron-scattering optical depth (Fig. 7, right) is significantly boosted by the presence of LMACHs, by about 0.01 overall, most of it due to the early stages of reionization. For the particular source efficiencies we have chosen here both optical depths fall within the 1σ interval given by the *WMAP* 5-year data.

These reionization histories are a direct consequence of the overall number of ionizing photons emitted by all active sources, shown in Fig. 8. In the case L3 where no source suppression occurs the number of photons emitted per timestep simply rises proportionally to the halo collapsed fraction, roughly exponentially. In contrast, in the fiducial self-regulated case L1 the initial exponential rise is halted around redshift $z \sim 16$ and rises very slowly (and moderately non-monotonically) until $z \sim 11$, at which point a sufficient number of (non-suppressible) HMACHs form to allow them to take over the evolution, while the LMACHs become highly suppressed. Therefore, similar to our earlier results in Iliev et al. (2007a) the late phase of reionization and overlap epoch, $z_{\rm ov}$, are dominated by HMACHs, while the LMACHs dominate the early phase of reionization and provide a significant boost to the electron-scattering optical depth, $\tau_{\rm es}$. Ultimately, by overlap in both simulations L1 and L3 there are 1.2–1.6 ionizing photon per atom emitted, slightly more in the self-regulated case due to its more extended reionization history which yields more recombinations per atom.

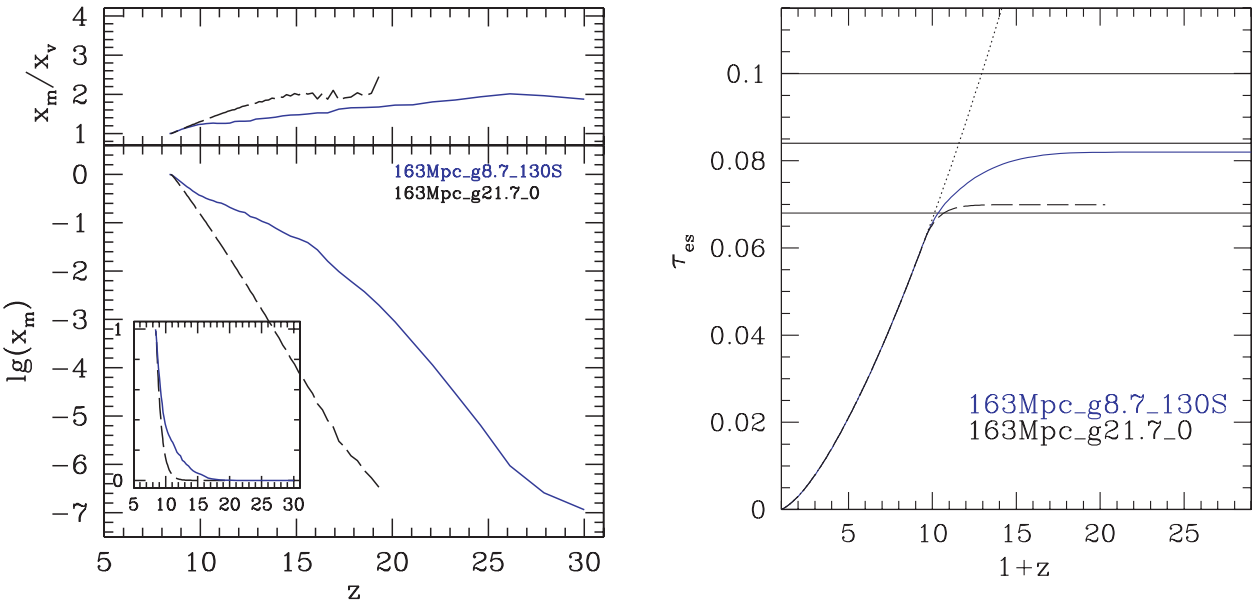

Figure 7. The effect of self-regulation on the reionization history and integrated electron-scattering optical depth. Left: mass-weighted reionization histories (bottom) and the ratio of the mean mass-weighted and volume ionized fractions, x_m/x_v (top) for our fiducial self-regulated case, 163Mpc_g8.7_130S (L1; blue, solid) and the corresponding non-self-regulated case with the same overlap epoch, 163Mpc_g21.7_0 (L3; black, long-dashed). The computational box size is $114/h = 163$ Mpc in both cases. The inset shows the same reionization histories in linear scale. Right: the corresponding electron-scattering optical depth, $\tau_{es}(z)$, integrated from redshift 0 to z for the same two simulations. The horizontal lines indicate the mean and 1σ band derived from the *WMAP* 5-year data, while the dotted line shows the value of τ_{es} for a fully ionized universe.

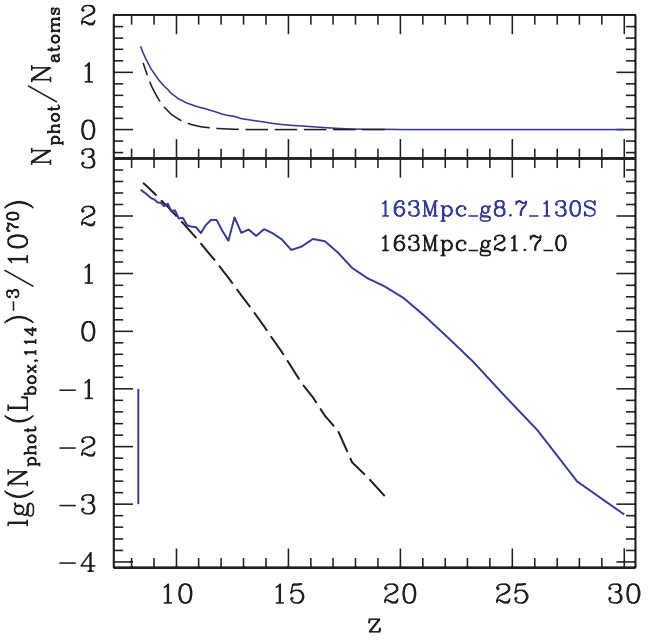

Figure 8. Bottom: number of ionizing photons emitted by all active sources in the computational volume per timestep; top: cumulative number of photons per total gas atom released into the IGM. Notation is the same as in Fig. 7. The vertical line marks the overlap redshift [$z(x_m = 0.99)$] for each case.

4.2 The effects of source efficiencies and box size

The reionization histories derived from our suite of simulations are shown in Fig. 9. The reionization history is monotonic in all cases, although due to the self-regulation the slope of the curves can vary significantly and in particular can become almost horizontal for short periods of time when the Jeans mass filtering compensates for the rise in source numbers. The exact redshifts at which certain reionization milestones, 10, 50 and 90 per cent by mass, are reached are listed in Table 2. We also list there the epochs when final overlap, which we define as the time when $x_m = 0.99$, i.e. at least 99 per cent of the mass is ionized, is reached in each case.

Our large-volume, self-regulated simulations L1 and L2 (Fig. 9, left-hand panels) have reionization histories which are very similar to each other, but offset by $\Delta z \sim 2$. Overlap is reached at $z = 8.3$ (6.7) in L1 (L2), corresponding to early (extended) reionization scenarios. The integrated electron-scattering optical depth for L1 is $\tau_{es} = 0.080$ for the early reionization case, well within the current *WMAP*5 1σ constraints. The corresponding value for the low-efficiency, extended reionization scenario L2 is 0.058, which is outside the 1σ range, but still within 2σ. On the other hand, the simulation volume (163 versus 53 Mpc) has little effect on the global reionization histories. This is in agreement with the results in Iliev et al. (2006a), which were derived by subdividing a $100\,h^{-1} = 143$ Mpc volume into smaller ones, which indicated that $\sim$20–30 h^{-1} Mpc box is sufficient to reliably derive the global mean reionization history. Most variations between the corresponding large- and small-box simulations result solely from the different random realizations in the two cases. At early times ($x_m < 0.01$) there are also departures due to cosmic variance – unlike the larger, 163 Mpc, volume the more limited 53 Mpc one does not contain any sources at $z > 25$ as those are statistically too rare to occur. Even when the very first haloes appear in the 53-Mpc volume, they are initially so few that they are subject to very high shot noise fluctuations. Once there are statistically significant numbers of sources in each size box, the reionization histories converge and any fluctuations thereafter are simply due to the different random realizations. There is also some effect from the higher resolution of the small-box simulations, due to the better resolved density field in those cases, which yields

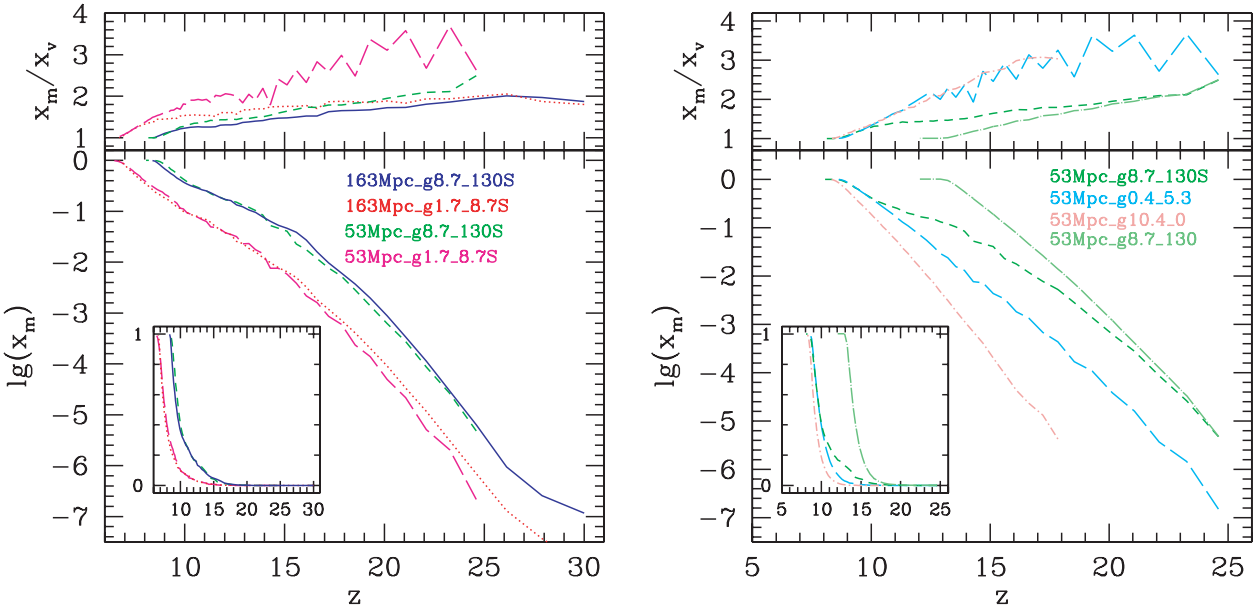

Figure 9. Bottom panels: mass-weighted reionization histories for our fiducial self-regulated cases (left) and with varying assumptions about the ionizing sources and their suppression (right). Top panels: the ratio of the mean mass-weighted and volume ionized fractions, x_m/x_v. All cases are labelled by colour and line type, as follows: (left) 163Mpc_g8.7_130S (L1; blue, solid), 53Mpc_g8.7_130S (S1; green, short-dashed), 163Mpc_g1.7_8.7S (L2; magenta, long-dashed) and 53Mpc_g1.7_8.7S (S2; red, dotted); (right) 53Mpc_g0.4_5.3 (S4; cyan, long dashed), 53Mpc_g10.4_0 (S5; light red, dot-short dashed) and 53Mpc_g8.7_130 (S3; light green, dot-long dashed). For ease of comparison, we show the fiducial case S1 on both plots.

slightly increased recombinations. This effect is rather minor here however.

In Fig. 9 (top left-hand panel) we show the ratio of the ionized fraction by mass, x_m and by volume, x_v, which is equal to the average density of the ionized regions in units of the mean (Iliev et al. 2006a). These ratios start at about 2 and remain above unity at all times, indicating that the reionization proceeds in an inside-out manner, with the high-density peaks (where the first sources preferentially form) being ionized on average earlier than the mean and low-density ones. On average the ionized regions are denser in the low-efficiency cases at any given redshift. This behaviour could be expected based on the typically smaller H II region sizes in those cases. They therefore stay in the immediate vicinity of the density peaks and do not propagate as much into the voids. The higher spatial resolution of the 53-Mpc cases also yields somewhat higher mean density of the ionized regions compared to the corresponding 163-Mpc box cases.

4.3 The effects of the source model: photon production efficiencies and minimum source mass

In Fig. 9 (right-hand panels) we show the corresponding reionization history results when the source models are varied. We also replotted one of our fiducial cases, S1, for facilitating direct comparison with the self-regulated cases. All reionization histories remain largely monotonic throughout the evolution, which therefore is a fairly robust feature, independent of the particular ionizing source properties assigned. However, a wide range of overlap epochs – from as early as $z = 12.9$ in the no-suppression case S3 to $z = 8.3$ in large-source-only case S5 – and a wide range of slopes of the reionization history evolution are observed.

Our fiducial case S1, along with S4, has the most extended reionization history of all, which starts with the formation of the first $10^8-M_\odot$ haloes at $z \sim 26$ and reaches overlap at $z = 8.9$. In comparison when all sources have the same efficiencies, but none are ever suppressed (case S3) the reionization history is very steep, roughly exponential, tracking the exponential rise of the collapsed fraction in haloes (cf. Fig. 3). The no-suppression, low-efficiency case S4 also produces a very extended history since it also starts with the formation of the first $10^8-M_\odot$ haloes and, by design, reaches overlap at roughly the same time as S1. However, the sources are necessarily much weaker in that case compared to S1 and S3, and therefore the ionized fraction starts much lower compared to the fiducial simulation and only catches up with the self-regulated case at late times ($x_m \gtrsim 0.25$). Finally, in the case of S5 only the HMACHs are active, and therefore the reionization starts late, at $z \sim 18$, but x_m rises exponentially, in proportion of the collapsed fraction in those massive haloes, reaching (again by design) overlap at the same time, $z = 8.3$ as the fiducial case. The lack of LMACHs therefore delays reionization considerably and naturally yields much lower integrated electron-scattering optical depth (0.071 compared to 0.084 for S1 and 0.078 for S4).

The mean overdensity of the ionized regions, x_m/x_v (Fig. 9, top right-hand panel) is above unity for all cases and at all times, demonstrating the robustness of the inside-out nature of reionization, in agreement with our original findings (Iliev et al. 2006a). Compared to our fiducial simulation, S1, the ratio x_m/x_v is significantly higher for cases S4 and S5. This is due to the fact that in the latter the H II regions are more tightly correlated with the density peaks, because the number of sources, which form at the density peaks, rises exponentially in these cases. Finally, simulation S3 shows an intermediate behaviour, similar to the fiducial case S1, but with somewhat faster decrease of the mean overdensity of the ionized regions, due to higher ionization of the low-density regions in this case.

The corresponding evolution curves of the cumulative number of ionizing photons emitted within each simulation are shown in Fig. 10. Indicated are also the number of photons which would have been emitted if no self-regulation had taken place (thin lines)

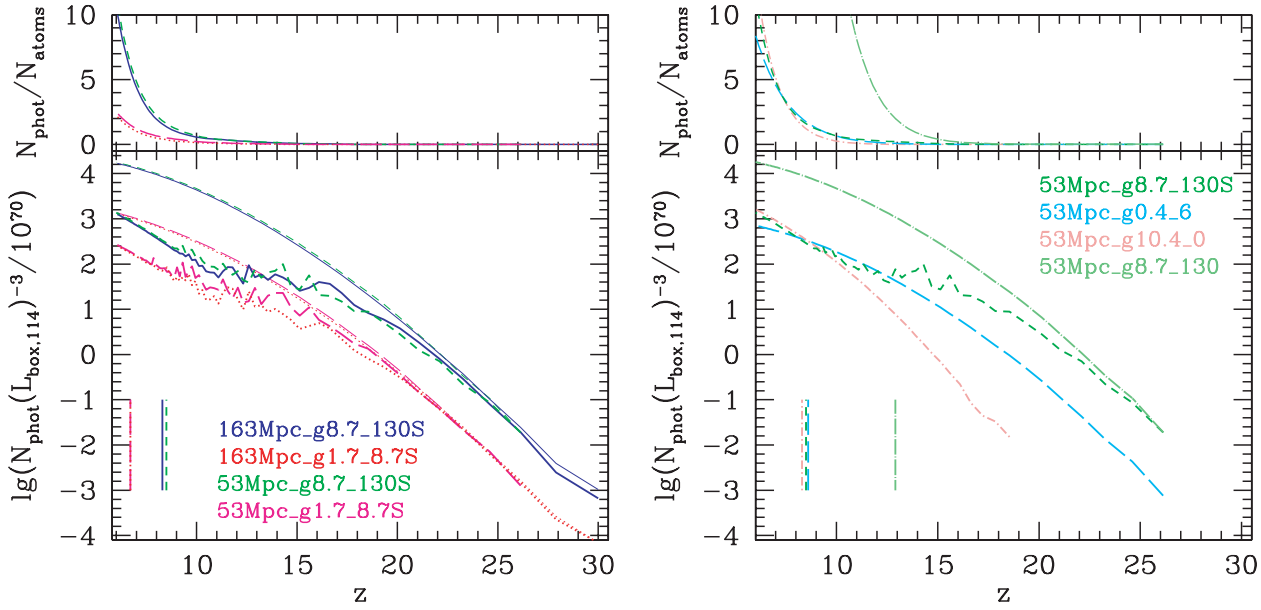

Figure 10. Bottom panels: the number of ionizing photons emitted by all sources (i.e. if there were no suppression; thin lines) and all active sources (thick lines) in the computational volume per timestep. Top panels: cumulative number of photons per total gas atom released into the IGM. Notation is the same as in Fig. 9. Vertical lines with same colours and line types mark the overlap redshift in each case.

and the overlap redshifts (vertical marks). Starting with our self-regulated cases (Fig. 10, left) we see that the LMACH suppression does not have a significant impact until $z \sim 20$–22, but after that has a great impact, reducing the overall number of emitted photons by up to a factor of $\sim$30 for highly efficient LMACHs and about a factor of 10 for less efficient ones. The number of ionizing photons emitted per timestep also becomes variable and non-monotonic function of redshift due to the complex interplay of suppression and new source formation during the self-regulation process. Eventually, by overlap (indicated for each case by the vertical lines) all LMACHs are suppressed and the number of photons continues to rise smoothly as ever more HMACHs form. The simulation box size makes little difference in the ionizing photon production, apart from modest variations due to the different random realization in each case. In terms of the cumulative numbers of emitted photons per atom (top panel), by redshift $z = 6$ up to 10 photons per atom are produced in the efficient-sources case and up to two photons per atom in the low-efficiency cases. However, at their respective overlap redshifts approximately the same numbers are produced, about 1.5, i.e. on average only about one recombination per every two atoms occurs during the evolution. Therefore, recombinations are relatively unimportant in these runs. The reason for this is that much of the density fluctuations are at very small scales, well below our radiative transfer grid resolution. This additional small-scale power can be added as subgrid clumping of the gas, calculated based on much higher resolution simulations. We will consider the effects of subgrid clumping in a companion paper (Koda et al., in preparation).

Turning our attention to the set of cases with different source efficiency models (Fig. 10, right), we first note that the three models which by construction have very similar overlap epochs (our fiducial case, S1, and cases S4 and S5) also have almost identical photon production numbers at overlap. All three reach this point in a very different manner, however. In run S4 all sources are quite weak, but no sources are ever suppressed, and the emissivity per timestep reaches the fiducial case once most LMACHs are suppressed in the latter case. In case S5, only the HMACHs are present and consequently its emissivity lags significantly at early times until eventually the exponential rise of those sources allows it to join the other two cases at $z \lesssim 11$. We also note that after overlap the photon emissivity in the no-suppression case S4 lags behind the others because its HMACHs are very inefficient and the collapsed fraction in LMACHs by this point does not rise as fast as the one for the HMACHs. Finally, the number of photons produced in the no-suppression, high source efficiency case, S3, simply follows the total collapsed fraction in all sources. It therefore initially rises almost exponentially, slowing down at later times as LMACHs become more common, at all times being roughly parallel to the curve for S4, eventually surpassing 10 photons per atom before $z = 10$. However, at its own (very early) overlap epoch, even this case produces the same number of photons as the others, about two per IGM atom.

In Figs 11 and 12 we show slices through our simulation volume showing the geometry of the H II regions at $x_m \sim 0.5$, overlaid on the corresponding density field for all our simulations. We also mark the cells containing active sources (blue/dark). Comparing first the large 163 box cases (Fig. 11), we note that in the fiducial L1 there are many more active sources compared to L2. In contrast, in simulation L3 there are many fewer sources due to its higher mass cut-off for sources. The large-scale structures are quite similar in size and shape in the two self-regulated simulations, but the fiducial case L1 yields much more small-scale ionized patches even though it reaches half-ionized state noticeably earlier, at which point there are many fewer, and more clustered, sources. The reason for this apparently counterintuitive behaviour is that the much weaker sources in case L2 have difficulties fully ionizing their own cells (which, at 445 kpc h^{-1} linear size, are relatively large), and therefore a large number of sources are needed to produce a sizeable fully ionized patch. In contrast, the much more efficient sources in our fiducial case L1 easily ionize their own cell, resulting in many small-size H II regions, instead of the more scattered, partially ionized cells in L2. On the other hand, in the non-self-regulated case L3

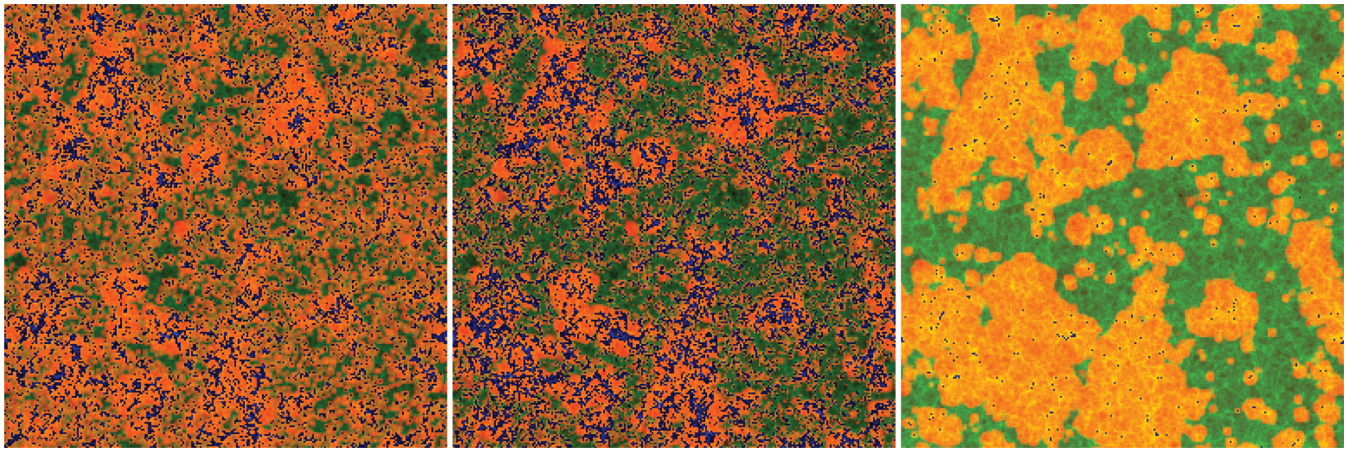

Figure 11. Spatial slices of the ionized and neutral gas density from our radiative transfer simulations with box size 163 Mpc: left, 163Mpc_g8.7_130S (L1); middle, 163Mpc_g1.7_8.7S (L2); and right, 163Mpc_g21.7_0 (L3), all at box-averaged ionized fraction by mass of $x_{\rm m} \sim 0.50$. Shown are the density field (green) overlaid with the ionized fraction (red/orange/yellow) and the cells containing active sources (dark/blue).

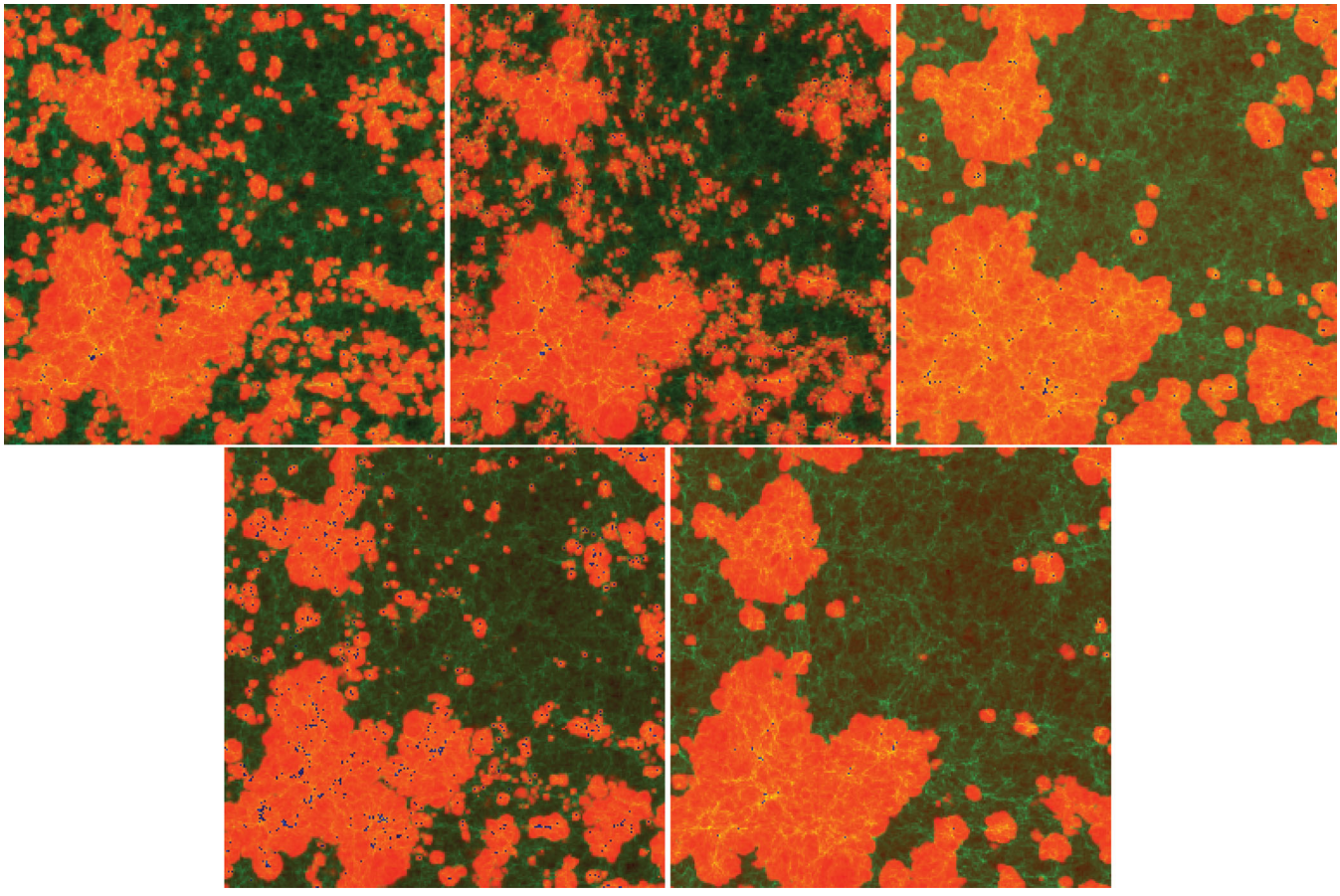

Figure 12. Spatial slices of the ionized and neutral gas density from our radiative transfer simulations with box size 53 Mpc, all at box-averaged ionized fraction by mass $x_{\rm m} \sim 0.50$. Shown are the density field (green) overlaid with the ionized fraction (red/orange/yellow) and the cells containing sources (dark/blue). Shown are (left to right and top to bottom) cases 53Mpc_g8.7_130S (S1), 53Mpc_g1.7_8.7S (S2), 53Mpc_g8.7_130 (S3), 53Mpc_g0.4_5.3 (S4) and 53Mpc_g10.4_0 (S5).

we find many fewer, larger ionized regions, in agreement with our previous results (Iliev et al. 2008a; Friedrich et al. 2011). The large-scale structures have some similarities to the ones found in the self-regulated cases, as could be expected given that all simulations share the underlying large-scale cosmic structures. However, the ionized regions are more merged together, and there is far less small-scale structure due to the absence of low-mass (LMACHs and smallest HMACHs) weaker sources. In that case, there are also no partially ionized regions, since there is no LMACHs suppression due to Jeans mass filtering (i.e. sources do not ever die), and all sources are sufficiently luminous to completely ionize their own local volume.

The corresponding images from our small-box simulations at the same ionized fraction of $x_m \sim 0.5$ are shown in Fig. 12. Once again, the large-scale structures, which tend to strongly correlate with the underlying distribution of density and clustered haloes, are generally quite similar. In contrast, ionized patches produced by multiple, less biased sources whose distribution does follow the knots and filaments of the Cosmic Web are much more irregularly shaped. There are significant differences in the smaller scale structures among the range of simulations. The self-regulated cases, S1 and S2, and to a lesser extent case S4 have the most small-scale structure, including both small H II regions and rough, irregularly shaped large H II region boundaries. In contrast, S5, which does not include the LMACHs, and S3, which includes efficient, unsuppressible LMACHs, both yield many fewer ionized patches with smoother boundaries, which reflects the rarity and highly clustered nature of their active sources. We have presented a more detailed discussion of the H II region geometry, size distribution and topological characteristics in a recent companion paper based on a subset of the current suite of simulations (Friedrich et al. 2011).

5 OBSERVATIONAL SIGNATURES

We now turn our attention to the reionization observables and specifically how are they related to the assumed source populations and their efficiencies. A better understanding of these dependencies should allow us in turn to use the observational data to constrain the properties of the reionization sources. Our main focus will be on the redshifted 21-cm signatures, although we also briefly discuss the photoionization rates in the IGM, related to the measurements of the Gunn–Peterson effect and the gas temperature.

5.1 Photoionization rates

An important, if indirect, observable signature of reionization is the mean photoionization rate in the IGM. At present, this quantity has only been measured for the post-reionization IGM at $z < 6$, derived based on the small residual neutral fraction and its corresponding Lyα optical depth. It therefore characterizes only the final EoR stages, around and after overlap.

The redshift evolution of the mean photoionization rates, Γ, averaged over our simulation volume for our large-box simulations L1–L3 is shown in Fig. 13. In overall curve shape and timing the evolution roughly mirrors the reionization histories for these three cases. This is natural, since the mean photoionization rate is the average of the fraction from the ionized regions, where the Γ values are high and fairly uniform at $\sim 10^{-12}\,\mathrm{s}^{-1}$ (see also Fig. 14), and the neutral regions, where $\Gamma \sim 0$. The mean photoionization rate in our fiducial case L1 initially rises very fast, until self-regulation suppressing the LMACHs becomes widespread at $z \sim 16$, at which point they stall at around $10^{-14}\,\mathrm{s}^{-1}$ until the HMACHs become sufficiently abundant to dominate the evolution, which occurs around $z \sim 11$. At this time the mean rates resume their steep rise, eventually reaching a peak of $\Gamma \sim$a few $\times\, 10^{-12}\,\mathrm{s}^{-1}$ by overlap. The situation is somewhat different for the low photon production efficiency case, L2. The mean Γ values are much lower in this case, by 2–4 orders of magnitude, and are mostly rising monotonically throughout the evolution, as the Jeans mass suppression effects are milder in this scenario. The peak value reached in this case is about $10^{-12}\,\mathrm{s}^{-1}$, in rough agreement with the measured one at $z \sim 6$. However, we should note that any direct comparisons to observationally derived values are at best approximate since our simulations at present do not take into account the Lyman-limit systems (LLS), which are likely to limit the growth of the mean-free path of the photons and thus limit Γ as well. Before overlap the mean-free path is dictated by the remaining neutral regions and the (still fairly high) residual neutral gas fraction within the ionized regions, and therefore the LLS are unimportant and do not affect our simulation results. The same is probably not true after overlap. We will study the effects of LLS in future work. For our current purposes, the lack of LLS means that we cannot yet make a firm conclusion that the low-efficiency, late-overlap case, L2, fits the observations better than case L1.

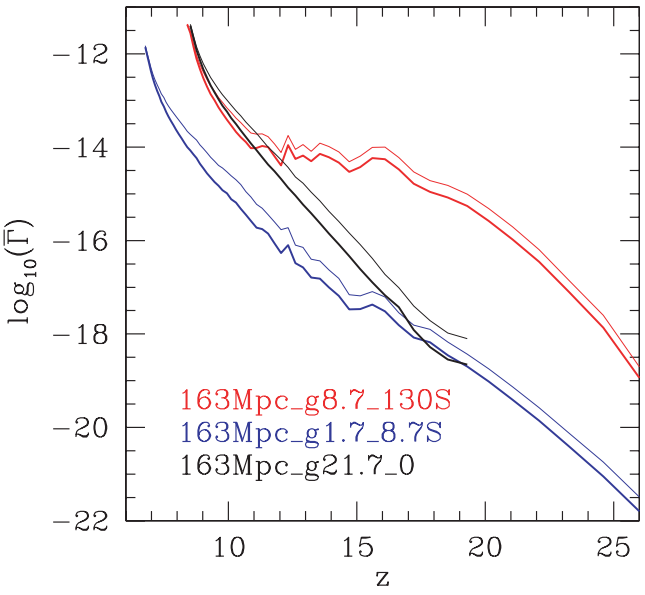

Figure 13. Evolution of the mean mass-weighted (thin lines) and volume-weighted (thick lines) photoionization rates in our computational volume for simulations 163Mpc_g8.7_130S (L1; red), 163Mpc_g1.7_8.7S (L2; blue) and 163Mpc_g21.7_0 (L3; black).

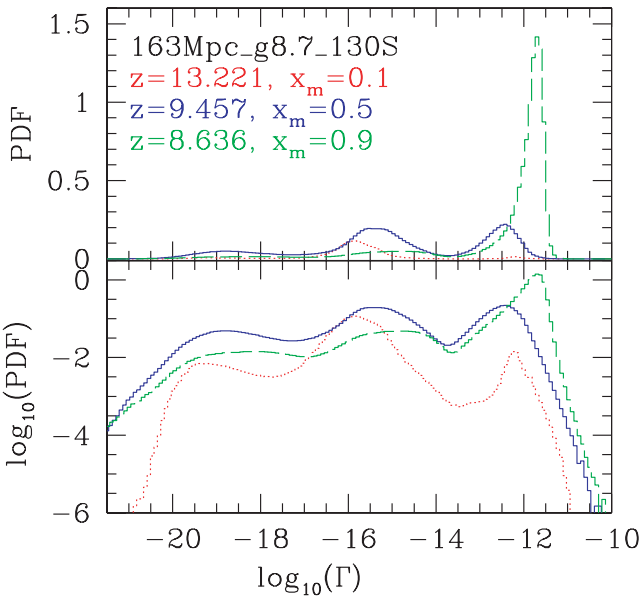

Figure 14. Photoionization rate PDFs for our fiducial case 163Mpc_g8.7_130S (L1) for epochs when the ionized fraction by mass is $x_m = 0.1$ (red), $x_m = 0.5$ (blue) and $x_m = 0.9$ (green) plotted in linear (top) and log (bottom) scales.

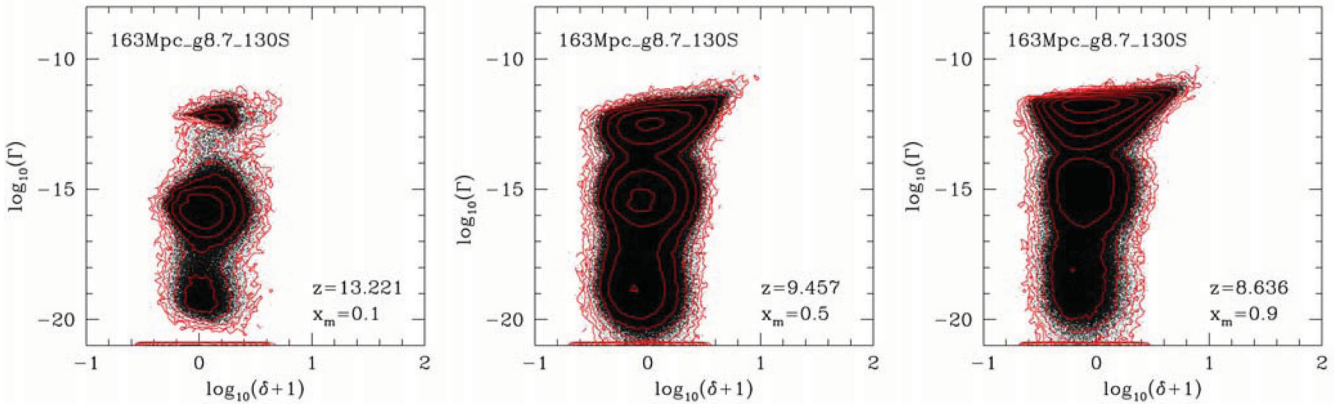

Figure 15. Cell-by-cell photoionization rate–overdensity correlation scatter plot for our fiducial case 163Mpc_g8.7_130S (L1) at mean ionized fractions by mass $x_{\rm m} = 0.1$ (left), $x_{\rm m} = 0.5$ (middle) and $x_{\rm m} = 0.9$ (right). Contours are logarithmic, from 10 cells up every 0.5 dex.

The mean photoionization rate for the non-self-regulated case L3 is intermediate between L1 and L2. It starts from very low values, $\sim 10^{-19}\,{\rm s}^{-1}$, when there are still only a very few HMACHs, around $z \sim 17$, but then rises sharply, roughly exponentially, and converges (by construction, since efficiencies were picked so that the two simulations reach overlap at the same time) to the values for L1 at later times.

The mass-weighted photoionization rates (thin lines) are significantly higher, by factors of up to 2–3, than the volume-weighted ones (thick lines) at all times and for all simulations. This is easy to understand given the inside-out nature of reionization, whereby the ionizing sources are found in dense regions, which boosts the mass-weighted means. Such large differences are interesting, however, since they can possibly skew the observationally derived values. Probes of the mean, low-density IGM will therefore yield considerably lower values for Γ than any measurements which are more sensitive to denser regions e.g. around sources.

Several illustrative probability distribution functions (PDFs; at cell size, here 445 kpc h^{-1}) of the photoionization rates for our fiducial simulation L1 are shown in Fig. 14. Plotted are the PDFs at early ($x_{\rm m} = 0.1$), intermediate ($x_{\rm m} = 0.5$) and late ($x_{\rm m} = 0.9$) stages of the evolution. These can be compared to the no LMACHs results we presented previously in Iliev et al. (2008b). At all times, the PDFs show a characteristic three-peaked profile. The rightmost peak, at $\Gamma_{-12} \sim 1$, is formed by the cells inside the H II regions, while the other two peaks correspond to partially ionized cells, predominantly located at the expanding I-fronts and in relic (i.e. recombining) H II regions. As we have shown in Iliev et al. (2008b), the ionization state is closely approaching or have reached equilibrium deep inside the ionized regions, but is far from equilibrium at the I-fronts. This holds true also for the current simulations. However, compared to our previous results, here the LMACHs suppression yields a significant fraction of volume occupied by relic H II regions and partially ionized cells and thus yields higher peaks at lower Γ values than that observed in the simulations without suppression.

In Fig. 15 we show scatter plots and the corresponding contour levels of the local photoionization rates, Γ, versus density in units of the mean, $1+\delta \equiv \rho_{\rm cell}/\bar{\rho}$ for our fiducial case L1 and $x_{\rm m} = 0.1$, 0.5 and 0.9. Overall, throughout most of the evolution there is a clear positive correlation between the density and the photoionization rate in the ionized regions. This could be expected, given that sources around which the photoionization rates peak form preferentially in high-density regions. However, the relationship between the two is complex, the correlation is rather weak and the scatter is significant. Similar to the PDFs discussed above, three peaks are observed. The high-Γ peak ($\Gamma \sim 10^{-12}\,{\rm s}^{-1}$) consists of the ionized cells, which are typically denser than average. The middle peak, at $\Gamma \sim 10^{-15}\,{\rm s}^{-1}$, corresponds to I-fronts and other partially ionized regions, while the cells with still lower values ($\Gamma \sim 0$) correspond to the neutral regions. At early times ($x_{\rm m} = 0.1$) the majority of cells are either neutral or partially ionized and the correlation with the local density is very weak, likely due to Poisson noise. When the process advances ($x_{\rm m} = 0.5$), a large population of fully ionized, high-Γ cells develops and within the H II regions the photoionization rate is fairly well correlated with the density, albeit still with a large scatter. On the other hand, the photoionization rates in the partially ionized regions show essentially no correlation with the density. At late times ($x_{\rm m} = 0.9$) these trends become even more pronounced and a quite tight correlation develops for the highest density regions. The overall behaviour is consistent with what we previously observed when no source suppression was present (Iliev et al. 2008b), albeit with some minor quantitative differences.

5.2 Redshifted 21-cm

The differential brightness temperature of the redshifted 21-cm emission with respect to the CMB is determined by the density of neutral hydrogen, $\rho_{\rm HI}$, and its spin temperature, $T_{\rm S}$. It is given by

$$\begin{aligned}\delta T_{\rm b} &= \frac{T_{\rm S} - T_{\rm CMB}}{1+z}(1-{\rm e}^{-\tau}) \\ &\approx \frac{T_{\rm S} - T_{\rm CMB}}{1+z}\frac{3\lambda_0^3 A_{10} T_* n_{\rm HI}(z)}{32\pi T_{\rm S} H(z)} \\ &= 28.5\,{\rm mK}\left(\frac{1+z}{10}\right)^{1/2}(1+\delta)\left(\frac{\Omega_{\rm b}}{0.042}\frac{h}{0.73}\right)\left(\frac{0.24}{\Omega_{\rm m}}\right)^{1/2}\end{aligned} \quad (3)$$

(Field 1959), where z is the redshift, $T_{\rm CMB}$ is the temperature of the CMB radiation at that redshift, τ is the corresponding 21-cm optical depth, assumed to be small when writing equation (3), $\lambda_0 = 21.16\,{\rm cm}$ is the rest-frame wavelength of the line, $A_{10} = 2.85 \times 10^{-15}\,{\rm s}^{-1}$ is the Einstein A-coefficient, $T_* = 0.068\,{\rm K}$ corresponds to the energy difference between the two levels, $1+\delta = n_{\rm HI}/\langle n_{\rm H}\rangle$ is the mean number density of neutral hydrogen in units

of the mean number density of hydrogen at redshift z,

$$\langle n_{\rm H}\rangle(z) = \frac{\Omega_{\rm b}\rho_{\rm crit,0}}{\mu_{\rm H} m_{\rm p}}(1+z)^3 = 1.909\times10^{-7}\,{\rm cm}^{-3}\left(\frac{\Omega_{\rm b}}{0.042}\right)(1+z)^3, \quad (4)$$

with $\mu_{\rm H} = 1.32$ being the corresponding mean molecular weight (assuming 24 per cent He abundance), and $H(z)$ is the redshift-dependent Hubble constant,

$$H(z) = H_0[\Omega_{\rm m}(1+z)^3 + \Omega_{\rm k}(1+z)^2 + \Omega_\Lambda]^{1/2} = H_0 E(z) \approx H_0\Omega_{\rm m}^{1/2}(1+z)^{3/2}, \quad (5)$$

where H_0 is its value at present, and the last approximation is valid for $z \gg 1$. Throughout this work, we assume that $T_{\rm S} \gg T_{\rm CMB}$, i.e. there is sufficient Lyα background to completely decouple the 21-cm transition from the CMB and that the neutral gas is heated well above the CMB temperature (due to e.g. a small amount of X-ray heating). Under these conditions the 21-cm line is seen in emission. These assumptions are generally well justified, except possibly at the earliest times (see e.g. Furlanetto, Oh & Briggs 2006b and references therein).

5.2.1 *Evolution of the patchiness*

In Fig. 16 we show position–redshift (position–frequency) slices cut through the simulated image cube as a radio array would see it (ignoring foregrounds). The spatial dimension is on the vertical, where we have duplicated the computational volume for visualization purposes, while the redshift/frequency is along the horizontal. Images are of the 21-cm emission differential brightness temperature signal extracted from our fiducial simulation L1, continuously interpolated in redshift/frequency including redshift-space distortions due to peculiar velocities. The volume is cut at an oblique angle in order to minimize artificial repetition of structures along any line of sight. The top and middle panels show images (in logarithmic, which shows better the residual H I fraction in the ionized regions and linear scale, which shows better the neutral structures) at the full simulation resolution, which is much higher than what current experiments will achieve given the sensitivity constraints. The bottom panel shows same data, but smoothed with a Gaussian beam and an integrated bandwidth which both roughly correspond to the values adopted in the LOFAR EoR experiment. To mimic the fact that an interferometer such as LOFAR is insensitive to the global signal, the mean signal at every frequency slice has been subtracted.

At high redshift, here $z > 11$, all H II regions are small and largely isolated. Smoothing the data to the LOFAR resolution (the ones for MWA and GMRT are even lower) renders such small structures undetectable. One needs at least $\sim$1 arcmin or better resolution for detecting them, making this regime a potential target for future, more sensitive experiments e.g. Square Kilometre Array (SKA). However, at intermediate redshifts (here $z \sim 10$) the ionized regions quickly grow by merging and remain clearly visible also after beamwidth and bandwidth smoothing. Even though some detail is lost, the large-scale structure of the ionization field remains visible all the way to the overlap epoch, here $z = 8.4$. As noted above, compared to the simulations with no self-regulation (e.g. Iliev et al. 2008b), the suppression of LMACHs introduces much more small-scale structure and many, mostly small partially ionized and relic H II regions. However, the smoothing to the radio array resolution largely eliminates this fine-scale structure and the result is, at least visually, not dramatically different from the case with no self-regulation. The minimum and maximum values of the differential temperature are also similar. We consider more quantitative measures of the 21-cm signal next.

5.2.2 *21-cm background: mean and rms*

The evolution of the mean differential brightness temperature and its rms fluctuations for our fiducial cases L1 and L3, which corresponds to our previous simulations with no self-regulation in Iliev et al. (2008b), are shown in Fig. 17. The presence of LMACHs and Jeans mass filtering yields initially a steeper decline of the mean 21-cm emission starting from $\nu \sim 80$ MHz ($z \sim 17$), at which point the LMACHs start to form in larger numbers (becoming 3σ haloes, cf. Fig. 3), while the HMACHs are still very rare. At $\nu \sim 130$ MHz ($z \sim 10$) the HMACHs in turn become 3σ haloes, i.e. relatively more common and the mean $\delta T_{\rm b}$ evolution for case L3 steepens, eventually reaching the same overlap epoch (by construction). In terms of detectability in experiments looking for rapid changes in the 21-cm signal as the Universe reionizes (Shaver et al. 1999; Bowman & Rogers 2010), this behaviour means that a self-regulated reionization scenario is even more difficult to detect than one which is not. While without self-regulation the global signal drops fast by about 25 mK between 130 and 150 MHz; with it the change at the higher frequencies is more gradual. The decrease with self-regulation is somewhat steeper at lower frequencies, $\nu = 80$–120 MHz, but it is still fairly gradual and thus will be difficult to detect.

Comparing the rms fluctuations averaged over LOFAR-like beam and bandwidth (Fig. 17, right) we see that the overall evolution follows similar paths in both cases. Early on very little of the gas is ionized and the fluctuations therefore simply follow the density ones. Only when a significant ionized fraction develops do the fluctuations depart from the underlying density. For simulation L3 this occurs fairly late, at $\nu > 110$ MHz, compared to much earlier, $\nu > 80$ MHz, for the fiducial simulation L1. At this point the rms fluctuations slightly dip, as the highest density peaks are ionized, which diminishes the mean $\delta T_{\rm b}$ but does not boost the fluctuations since the H II regions are still smaller than the smoothing size. As the H II regions grow, the fluctuations increase again, reaching a peak before the signal dips and eventually disappears as the IGM becomes highly ionized. The peak position (at 142 MHz) remains the same in both cases and is thus not affected by self-regulation. The rms fluctuations are lower with self-regulation, by about 1/3 at the peak. The reason for this is the lower mean differential brightness temperature in that case, as can be seen in the right-hand panel. When the fluctuations are normalized by the mean they become identical in the two cases once they surpass the density fluctuations ($\nu > 127$ MHz). Before that point, the fluctuations in case L3 closely follow the density ones, while in case L1 they are lower because the highest density peaks have already been ionized.

In Fig. 18 we show the evolution of the mean redshifted 21-cm differential brightness temperature for high versus low source efficiencies and different box sizes (left) and for varying source models (right). The lower photon efficiencies (simulation L2) predictably yield a more gradual transition of the global IGM from neutral to ionized state compared to our fiducial case L1. For example, the evolution from 25 to $\sim$0 mK occurs over $\sim$50 MHz, from 130 to 180 MHz. Such an evolution would make detection of the 'global step' even more difficult compared to the fiducial case. The computational volume adopted for the simulation makes very little difference to the

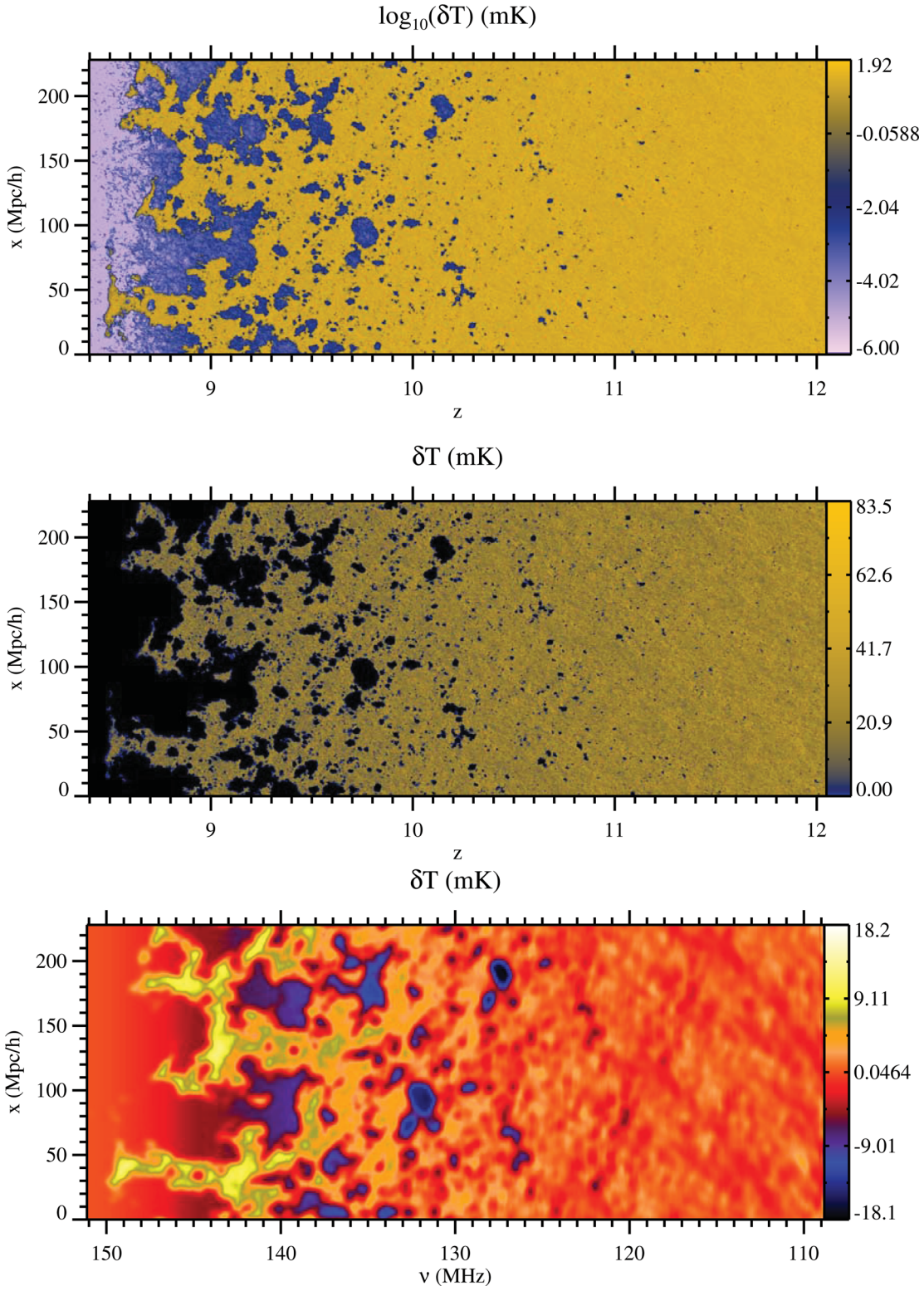

Figure 16. Position–redshift and position–frequency slices from our fiducial simulation 163Mpc_g8.7_130S (L1). These slices illustrate the large-scale geometry of reionization and the significant local variations in reionization history as seen at the redshifted 21-cm line. Observationally, they correspond to slices through an image-frequency volume of a radio array. The top and middle images show the differential brightness temperature at the full grid resolution in decimal log and linear scale, respectively. The bottom image shows the same δT_b data, but smoothed with a Gaussian beam of 3 arcmin and (top-hat) bandwidth of 0.45 MHz, roughly corresponding to the expected parameters for the LOFAR EoR observations. In order to mimic the behaviour of an interferometer, the mean signal has been subtracted for every frequency slice. The spatial scale is given in comoving Mpc and we note that for visualization purposes we have doubled (periodically) the box size in the spatial direction. The redshift-space distortions due to the peculiar velocities are also included.

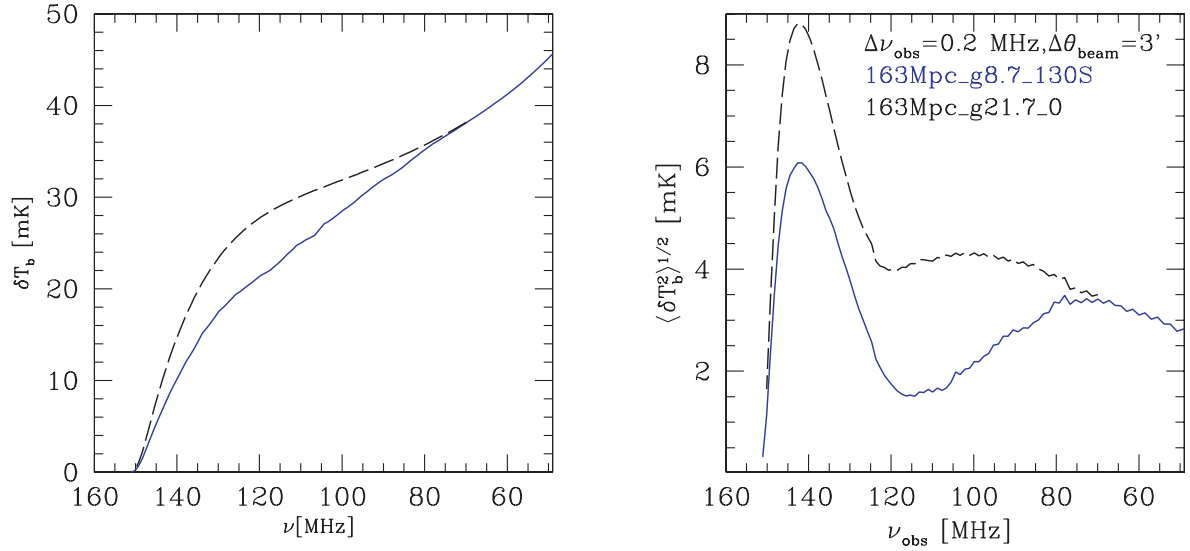

Figure 17. The evolution of the mean 21-cm background (left) and its rms fluctuations for Gaussian beamsize 3 arcmin and bandwidth 0.2 MHz and bandwidth frequency filter (right) versus observed 21-cm frequency. Shown are simulations 163Mpc_g8.7_130S (L1; blue, solid) and 163Mpc_g21.7_0 (L3; black, long-dashed).

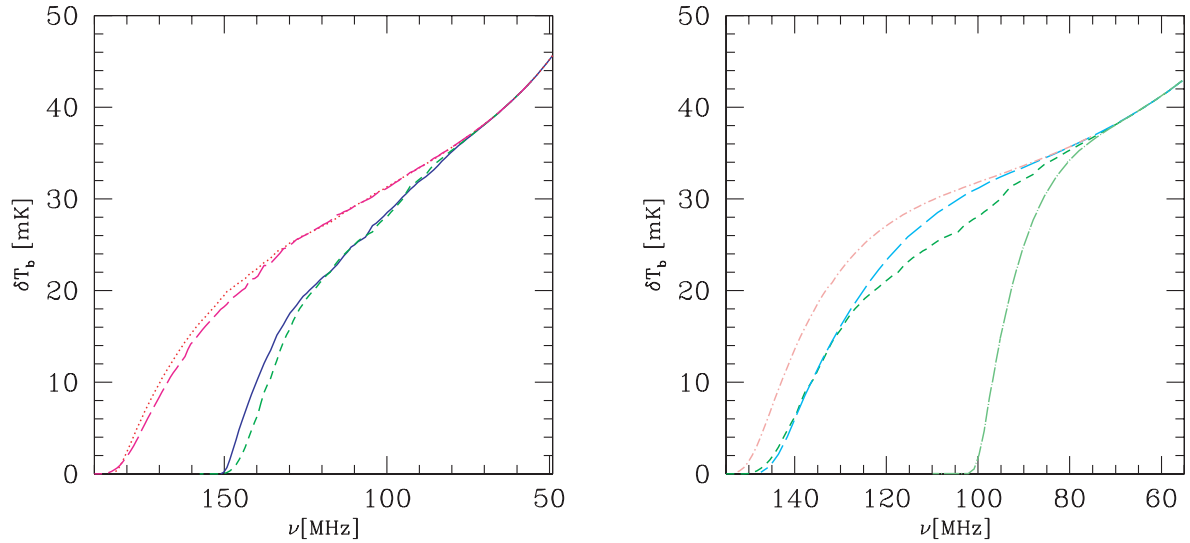

Figure 18. The evolution of the mean 21-cm background for our fiducial cases (left) and varying the source model (right). All cases are labelled by colour and line type, as follows: (left) 163Mpc_g8.7_130S (L1; blue, solid), 53Mpc_g8.7_130S (S1; green, short-dashed), 163Mpc_g1.7_8.7S (L2; magenta, long-dashed) and 53Mpc_g1.7_8.7S (S2; red, dotted); (right) 53Mpc_g8.7_130S (S1; green, short-dashed, same as in the left-hand panel), 53Mpc_g0.4_5.3 (S4; cyan, long dashed), 53Mpc_g10.4_0 (S5; light red, dot-short dashed) and 53Mpc_g8.7_130 (S3; light green, dot-long dashed).

predicted mean 21-cm signal, demonstrating again that the $37/h = 53$ Mpc box is sufficiently large to faithfully represent the mean reionization history. Most cases with varying UV source models and LMACHs suppression (Fig. 18, right) yield mean 21-cm histories which are quite similar to each other, a consequence of their analogous reionization histories. The only noticeable differences are at intermediate frequencies, between 90 and 130 MHz, where case S4 gives higher δT_b by up to 5 mK. The only significantly different evolutions are provided by cases S5 and S3. Those scenarios exhibit a sharper 21-cm step due to their faster, exponential rise of the ionizing photon emissivity, a consequence of the weak or no suppression in those cases.

The evolution of the rms 21-cm emission fluctuations for LOFAR-like beam and bandwidths corresponding to the same sets of simulations as in Fig. 18 are shown in Figs 19 (high versus low ionizing efficiencies and varying box size) and 20 (different UV source models). We show the same data versus observed frequency (left) and versus ionized fraction (right). The latter case takes away the reionization timing and allows comparison at the same stages of each reionization history.

In all cases the rms evolution roughly follows the same overall pattern, with an initial rise tracking the underlying density fluctuations when the IGM is mostly neutral, with a subsequent decrease when the first H II regions appear followed by a second, higher peak of the fluctuations at later times when the initially small ionizing regions grow, overlap locally and as a result match better the interferometer beam and bandwidth resolution, and a final decline when most hydrogen is ionized. However, despite this recurring pattern,

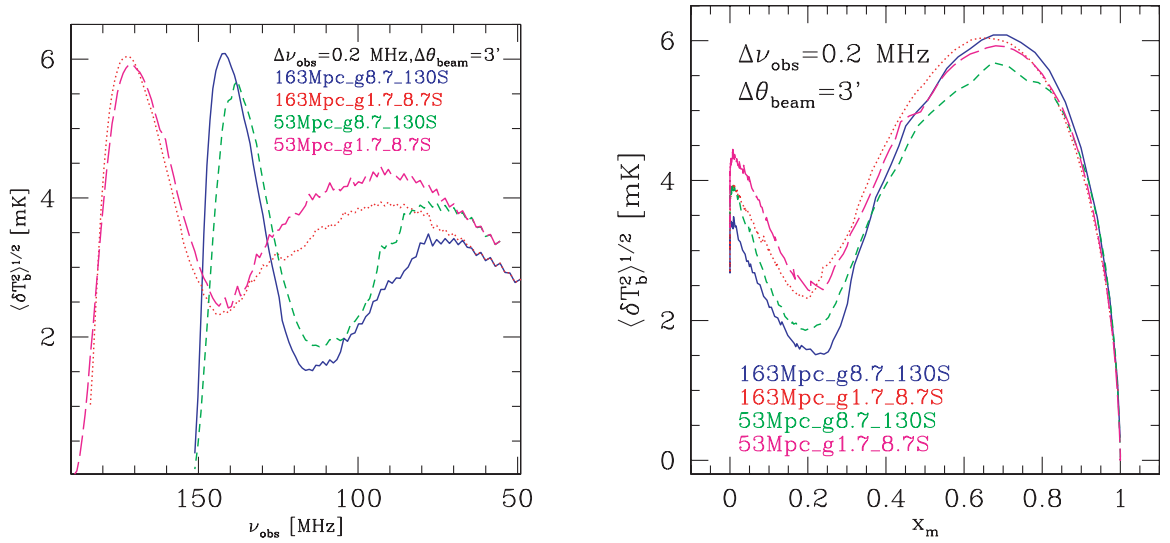

Figure 19. The evolution of the rms fluctuations of the 21-cm background, for beamsize 3 arcmin and bandwidth 0.2 MHz and bandwidth filter versus redshift (left) and versus mean mass-weighted ionized fraction (right). Shown are our fiducial simulations 163Mpc_g8.7_130S (L1; blue, solid), 53Mpc_g8.7_130S (S1; green, short-dashed), 163Mpc_g1.7_8.7S (L2; magenta, long-dashed) and 53Mpc_g1.7_8.7S (S2; red, dotted).

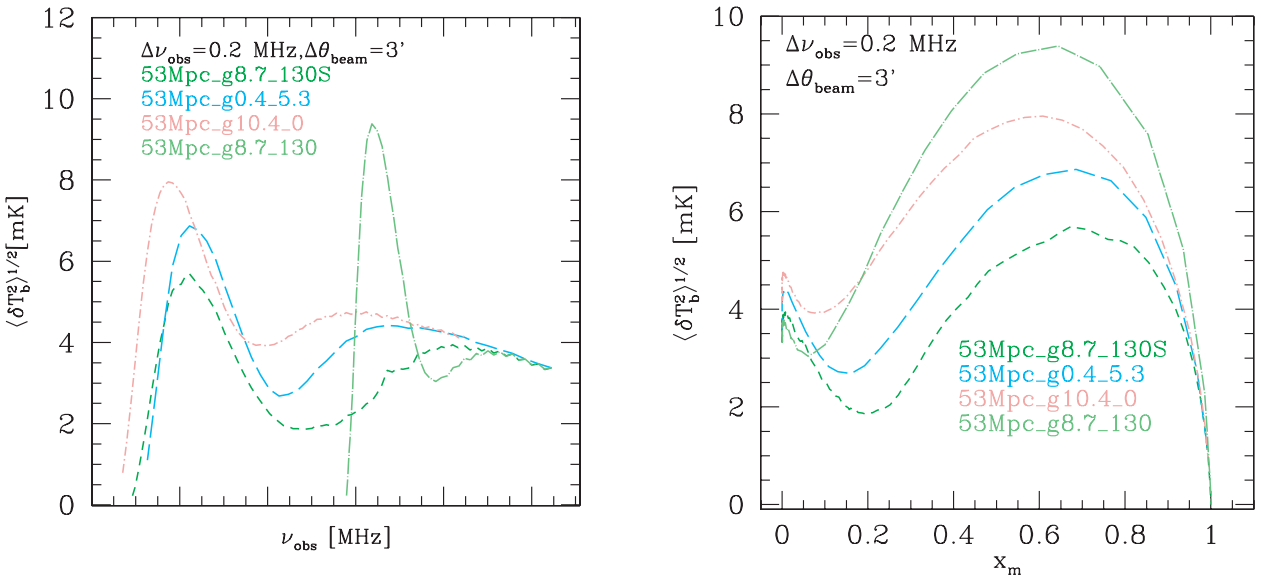

Figure 20. Same as Fig. 19, but for varying UV source models: 53Mpc_g8.7_130S (S1; green, short-dashed), 53Mpc_g0.4_5.3 (S4; cyan, long dashed), 53Mpc_g10.4_0 (S5; light red, dot-short dashed) and 53Mpc_g8.7_130 (S3; light green, dot-long dashed).

there are significant and often instructive differences among the models. Varying the ionizing photon emission efficiencies primarily changes the timing of the peak of the fluctuations (Fig. 19, left), from 142 MHz ($z = 9$) for the fiducial case L1 to 172 MHz ($z = 7.24$) for the low-efficiency case L2. However, as seen in Fig. 19 (right) this shift to later times is fully explained by the delayed reionization in the low-efficiency model and both curves peak at a mass-weighted ionized fraction $x_m \sim 0.7$. We note that the latter value is dependent, apart from the reionization parameters, also on the beam and bandwidth considered, as the peak is reached when the typical H II region size is best matched to the radio array resolution; therefore, the peak occurs earlier in the reionization history for higher resolution observations. The fluctuations dip due to the earliest H II regions also occurs at the same point of the reionization history ($x_m \sim 0.2$–0.25) in both cases, but the lowest rms values reached differ significantly at 1.5 mK for the fiducial model versus 2.3 mK for the low-efficiency one. The reason for this is that in the former case the bottom occurs earlier, when the sources responsible are rarer, more biased and therefore ionize the highest density peaks, which results in a larger decrease in the fluctuations. The simulation volume (and, correspondingly, radiative transfer grid resolution) has moderate, but appreciable effect on the 21-cm fluctuations in our fiducial case. The peak height is decreased by 7 per cent, from 6.1 to 5.7 mK, but it is also shifted to earlier time/lower frequency (to 138 MHz). Interestingly, while the dip of the fluctuations is also shifted to lower frequency for the smaller box simulation, the lowest rms value is in fact higher. At first sight, this appears counterintuitive, since naively we might expect that the higher grid resolution in the smaller volume to yield a larger rms decrease (since the density peaks where the first sources form are resolved better in this case).

What actually occurs here is more complicated, however. Statistically, there are fewer (and lower) high-density peaks in the smaller volume, which diminishes the effect of the very first sources on the fluctuations. Furthermore, the very first sources form later in the smaller box, again due to its much smaller volume, which means that the 21-cm rms fluctuations track the density ones for somewhat longer. On the other hand, the effects of box size and resolution for the low-efficiency cases are small and manifest themselves solely through the higher underlying density fluctuations in the small-box, high-resolution simulation. Both the peak and dip reach the same value for the two cases and occur at the same frequencies.

As shown in Fig. 20, variations of the ionizing source model yield a wider variety of 21-cm rms evolutions. Interestingly, all models exhibit the basic evolution features seen in our fiducial simulations – the initial dip of the rms value when the first H II regions appear, followed by a (relatively narrower) peak at later times when the process is sufficiently advanced for the typical patch size to roughly match the radio beam and bandwidth (though we note that more extreme, and unrealistic, source models can produce rms fluctuations with a very different shape, see Appendix A). There are significant variations in the details of the evolution, however. The simulations with no LMACHs suppression but same overlap as in our fiducial case (S4 and S5) yield rms fluctuation peaks which are at roughly the same frequency as the fiducial case ($\nu \sim 140$ MHz, more specifically), but the $\langle \delta T_{\rm b}^2 \rangle^{1/2}$ peak values are up to 50 per cent higher (∼7–8 mK). The early reionization, no-suppression case, S3, gives a still higher peak value, reaching almost 10 mK, and a narrower peak.

When plotted against the mass-weighted ionization fraction (Fig. 20) there is only a modest variation in the reionization stage (i.e. ionized fraction, $x_{\rm m}$) at which the rms peak is reached: $x_{\rm m} \sim 0.6$–0.7 (i.e. relatively late in the reionization history). However, as we noted above, the peak value itself varies significantly between the simulations. It is highest for the high-efficiency no-suppression model, S2 ($\langle \delta T_{\rm b}^2 \rangle^{1/2} = 9.4$ mK compared to 3.8 mK for our fiducial case S1). Lower efficiency and no-suppression (S4) results in a significantly lower peak (6.9 mK), while when only HMACHs are present (S5) the peak is again quite high ($\langle \delta T_{\rm b}^2 \rangle^{1/2} = 8.0$ mK). Therefore, in general more abrupt reionization scenarios (S3 and S5) result in higher fluctuations at the peak, while more extended ones (due to self-regulation or lower efficiencies) give lower rms peak values.

5.2.3 *21-cm background fluctuations: power spectra*

We now turn our attention to the (3D) power spectra of the 21-cm emission derived from our simulations. We construct the brightness temperature data cube in the redshift space using what we term the PPM-RRM scheme (Mao et al. 2012), as follows. We first develop an adaptive-kernel, SPH-like approach to compute the bulk-flow velocity of the IGM at any position, directly from N-body particle data. We paint the particle mass by the hydrogenic neutral fraction of the radiative transfer grid cell in which the particle resides. Then N-body particles are Doppler-shifted to their apparent locations by line-of-sight bulk-flow velocity, new smoothing kernel lengths are computed using the new particle positions in redshift space, and halo-excluded particle data (i.e. H I mass) are again smoothed on to a regular, redshift-space grid at radiative transfer grid resolution. Then we compute the redshift-space H I density fluctuation and the 21-cm brightness temperature measured in redshift space by

$$\delta T_{\rm b}^{\rm s}(\boldsymbol{s}) = \widehat{\delta T}_{\rm b}(z_{\rm cos}) \left[1 + \delta^{\rm s}_{\rho_{\rm HI}}(\boldsymbol{s})\right], \tag{6}$$

where $\widehat{\delta T}_{\rm b}$ was defined like $\delta T_{\rm b}$ in equation (3), but with $1 + \delta$ replaced by its mean value $\bar{x}_{\rm HI} \equiv \langle n_{\rm HI} \rangle / \langle n_{\rm H} \rangle$, i.e. the mean neutral fraction. We compute the redshift-space power spectrum using fast Fourier transform (FFT). We refer the reader to Mao et al. (2012) for a detailed discussion of this methodology.

In Fig. 21 we compare the 21-cm power spectra for models L3 (equivalent to our previous simulations with no self-regulation) and our fiducial self-regulated case L1 for several representative stages of reionization. In both cases the 21-cm power spectra are initially close to a power law, with no characteristic scale, and with only the L3 power flattening out at small scales due to its lack of very small-scale structures and smooth, large H II regions. Overall, there is always less power in our fiducial case, by a factor of 50 per cent (in Δ_k), or more. As ionized regions continue to expand, a characteristic scale starts to emerge, which for these particular simulations is around wavenumbers $k \sim 0.2$–$0.8\,h\,{\rm Mpc}^{-1}$. Interestingly, this feature shows up at the same scales regardless of the presence of

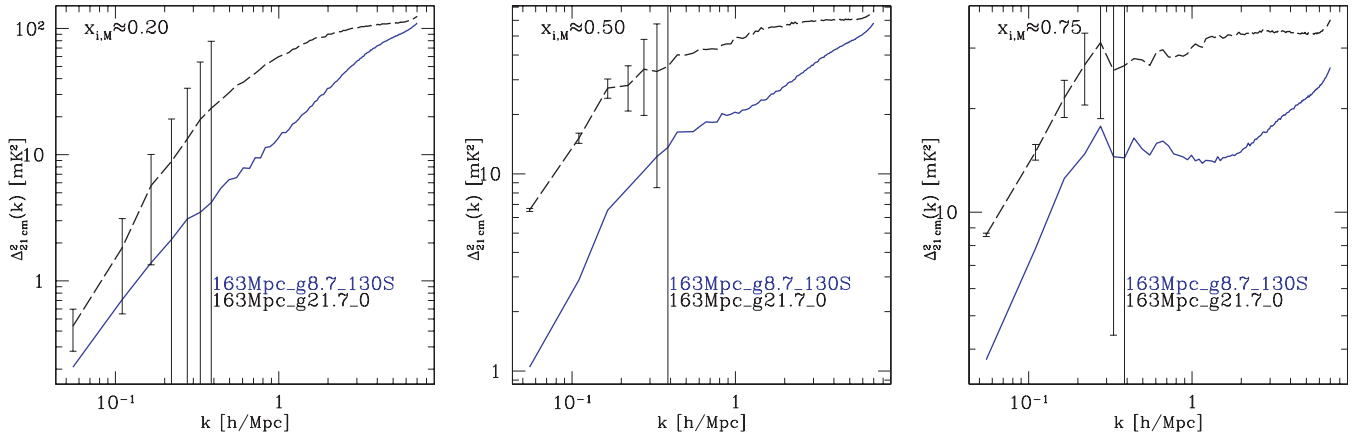

Figure 21. The effect of self-regulation on the 21-cm differential brightness temperature fluctuation power spectra. Shown are the epochs at which the ionized fractions are (left) $x_{\rm m} = 0.2$, (middle) $x_{\rm m} = 0.5$ and (right) $x_{\rm m} = 0.75$ for our fiducial self-regulated case, 163Mpc_g8.7_130S (L1; blue, solid) and the corresponding non-self-regulated case with the same overlap epoch, 163Mpc_g21.7_0 (L3; black, long-dashed). The error bars are thermal noise of the 21-cm power spectra measurement by the MWA experiment.

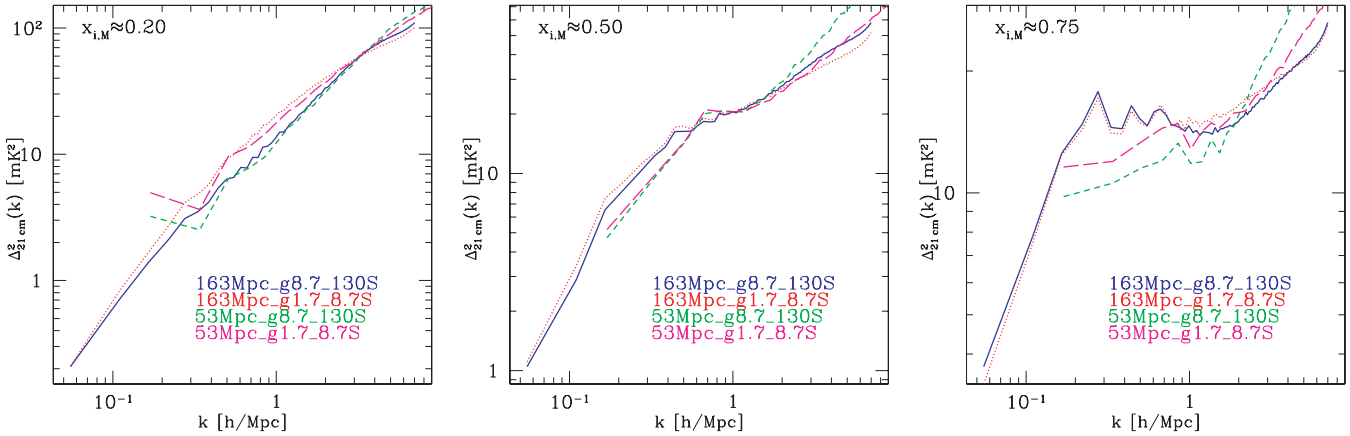

Figure 22. The 21-cm differential brightness temperature fluctuation power spectra. Shown are the epochs at which the ionized fractions are (left) $x_m = 0.2$, (middle) $x_m = 0.5$ and (right) $x_m = 0.75$. All cases are labelled by colour and line type, as follows: 163Mpc_g8.7_130S (L1; blue, solid), 53Mpc_g8.7_130S (S1; green, short-dashed), 163Mpc_g1.7_8.7S (L2; magenta, long-dashed) and 53Mpc_g1.7_8.7S (S2; red, dotted). Error bars are for the MWA experiment with 512 antennas distributed within a radius of 750 m, with a detection in a 1000-h integration time using a single station beam with an 8-MHz bandwidth.

LMACHs, indicating that the characteristic H II region size is caused by the clustering of the unsuppressible HMACHs.

The upcoming first-generation of 21-cm interferometer arrays will measure the power spectra of 21-cm brightness temperature fluctuations. If the foreground contamination can be removed cleanly (see Liu & Tegmark 2012 and references therein), the dominant source of error is the thermal noise for these observations. We model the thermal noise in the standard approach (e.g. McQuinn et al. 2006), and specifically estimate it for the MWA experiment which exemplifies the first-generation arrays. We assume a configuration of 512 antennas distributed within a radius of 750 m, with a detection in 1000-h integration time using a single station beam with an 8-MHz bandwidth. System specifications such as system temperature and effective collection area are assumed in the same way as in Bowman, Morales & Hewitt (2007). In Fig. 21, we plot the measurement errors from the MWA on top of the 21-cm power spectra signal for the case L3. At the late time of reionization ($x_i \geq$ 50 per cent), the 21-cm power spectra measurement by the MWA experiment can distinguish the case of reionization by HMACHs alone (case L3) from that by both HMACHs and LMACHs (case L1), at small wavenumbers/large scales, $k \lesssim 0.2\,h\,\mathrm{Mpc}^{-1}$.

On the other hand, the assumed ionizing source efficiencies have only a minor effect on the power spectra once the self-regulation is included (Fig. 22). The characteristic H II region scale is the same in the two cases and arises at the same point in the reionization history. The modest differences in the power spectra at the early stages of reionization arise due to the preferential ionization of the high-density peaks (where the first sources form). At the same ionized fraction x_m there are many more sources in the low-efficiency case, forming very small H II regions, versus fewer, larger ones in the high-efficiency case. As a result, the low-efficiency model yields less power at very small scales ($k > 4$), but a little more power at intermediate scales ($k = 0.2$–2). At the middle and late stages of reionization the two power spectra at the same x_m are largely identical, with only small differences due to the different amount of small-scale structure. The simulation volume utilized also has little effect on the derived power spectra at early times, essentially just shifting the range of k over which the results obtained are reliable. However, as the H II regions grow at intermediate and late stages of reionization, their sizes become comparable to the simulation volume and as a result the $37/h = 53$ Mpc volume cannot represent the bubble size distribution correctly and the calculated power spectra completely miss the characteristic H II region scale. On the other hand, the smaller volumes yield more fluctuation power at small scales, due to their superior spatial grid resolution. These results argue for a strong caution when using small (sub-$100/h = 143$ Mpc) boxes for predicting any EoR observables at late times.

Turning our attention to the varying source models (Fig. 23), several trends become clear. The lack of Jeans mass filtering (case S3 versus the fiducial S1) results in much more power at large scales ($k < 5\,h\,\mathrm{Mpc}^{-1}$ early, $k < 2\,h\,\mathrm{Mpc}^{-1}$ at late times), but considerably less power on small scales, a consequence of the large H II regions with smooth boundaries in case S3 produced by its luminous and strongly clustered sources. This also results in very flat power spectra for S3, with roughly constant power at all scales at intermediate and late times ($x_m = 0.5$ and 0.9). On the other hand, if the LMACHs are not present at all (case S5 versus S3) there is more power on all scales during the early stages of reionization ($x_m = 0.1$). However, at intermediate and late stages of reionization, the power spectra for the large-source-only case S5 remain steeper, with considerably more power at small scales, which results in cases S3 and S5 power spectra crossing at $k \sim 2\,h\,\mathrm{Mpc}^{-1}$. The reason for this somewhat counterintuitive behaviour is that the same reionization stage is reached in S5 much later than in S3, by which time there are many more and thus less clustered HMACHs, which in turn results in more power at small scales. Finally, the low-efficiency case with no self-regulation (S4) yields very similar power spectra to our fiducial simulation S1 throughout the evolution, with only slightly more power at large scales during the early and intermediate stages, and slightly less small-scale power at the late stages. This suggests that the S1 and S4 scenarios might be difficult to distinguish solely through power spectra measurements.

5.2.4 21-cm background fluctuations: PDFs and non-Gaussianity

The 21-cm power spectra would fully characterize the emission field if the differential brightness distribution were purely Gaussian. However, generally that is not the case during reionization, as we have shown in previous work (Mellema et al. 2006b; Iliev et al. 2008a; Harker et al. 2009). The PDFs of δT_b could be significantly non-Gaussian, particularly at the later stages of reionization (Mellema et al. 2006b) and their measured skewness can be used to

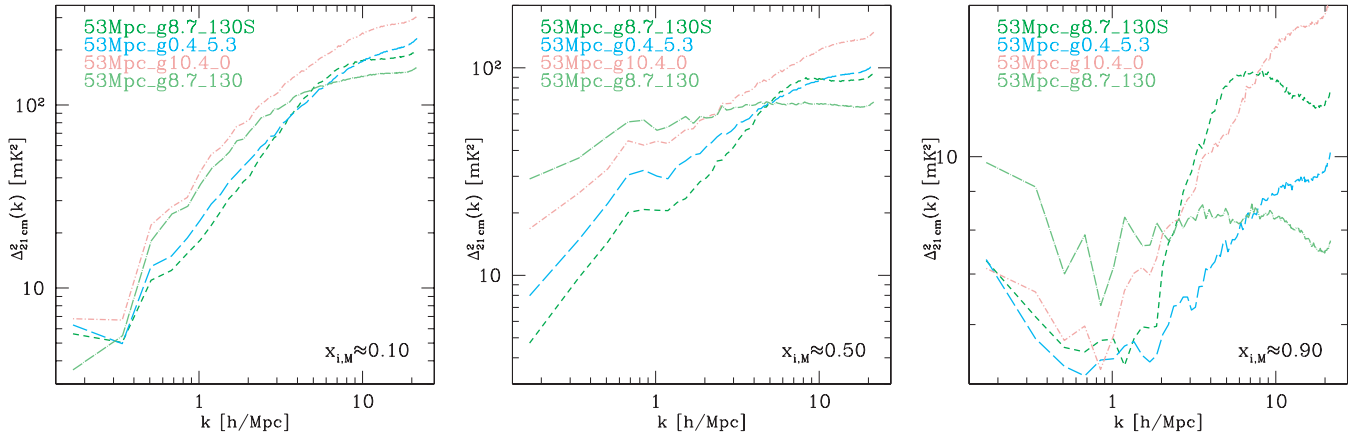

Figure 23. The 21-cm differential brightness temperature fluctuation power spectra for varying source models. Shown are the epochs at which the ionized fractions are (left) $x_m = 0.1$, (middle) $x_m = 0.5$ and (right) $x_m = 0.9$. All cases are labelled by colour and line type, as follows: 53Mpc_g8.7_130S (S1; green, short-dashed), 53Mpc_g0.4_5.3 (S4; cyan, long dashed), 53Mpc_g10.4_0 (S5; light red, dot-short dashed) and 53Mpc_g8.7_130 (S3; light green, dot-long dashed).

discriminate between different reionization scenarios (Harker et al. 2009). The PDFs and their evolution could also be used to derive the reionization history of the IGM (Gluscevic & Barkana 2010; Ichikawa et al. 2010).

The 21-cm cell-by-cell PDFs with and without the presence of suppressible LMACHs at three representative stages of reionization ($x_m = 0.1$, 0.5 and 0.9) are shown in Fig. 24. Early on ($x_m = 0.1$) the distributions are mostly following the underlying density field, and as a consequence are mostly Gaussian. There is a non-Gaussian tail for high differential brightness temperatures due to density non-linearities. Reionization itself introduces some non-Gaussianity at low δT_b as the first H II regions form around the highest density peaks, moving the corresponding cells into the extreme left of the distributions (i.e. holes in the neutral gas distribution, with $\delta T_b \sim 0$). This slightly skews the distribution towards below-average (i.e. negative in $\delta T_b - \delta\bar{T}_b$) temperature values since the low-density regions remain more neutral, on average. Self-regulation mitigates those trends somewhat, as the LMACHs are less clustered and more uniformly distributed throughout the volume, rather than being only at the highest density peaks (which are strongly clustered, as a consequence of the Gaussian density field statistics, see Fig. 6). For the same reasons, the PDF with self-regulation is also slightly less wide than without, as evidenced by the Gaussian distributions with the same mean and width as the actual PDFs. During the later stages of reionization, the distribution becomes ever more non-Gaussian, with the most prominent feature due to the ionized regions ($\delta T_b - \delta\bar{T}_b < 0$). The remaining neutral regions are a mixture of voids (low, but positive $\delta T_b - \delta\bar{T}_b$) and a few remaining higher density regions (e.g. filaments) which form the 21-cm bright non-Gaussian tail. Once again, both features are much reduced when the LMACHs suppression is taken into account due to the weaker clustering of these sources. The width of the PDFs decreases over time, and does so faster when the LMACHs are not present. At late times the two distributions have almost the same means and widths, although the actual distributions remain significantly different.

When the same PDFs are smoothed with a 5-Mpc h^{-1} window (roughly similar in size to e.g. the LOFAR beam, albeit here we use a different window shape for simplicity of the calculations), the results become notably different (Fig. 25). The smoothed PDFs at early times ($x_m = 0.1$) become significantly more Gaussian, although some residual non-Gaussian tails remain at both high and low δT_b. As could be expected, the smoothed distributions also

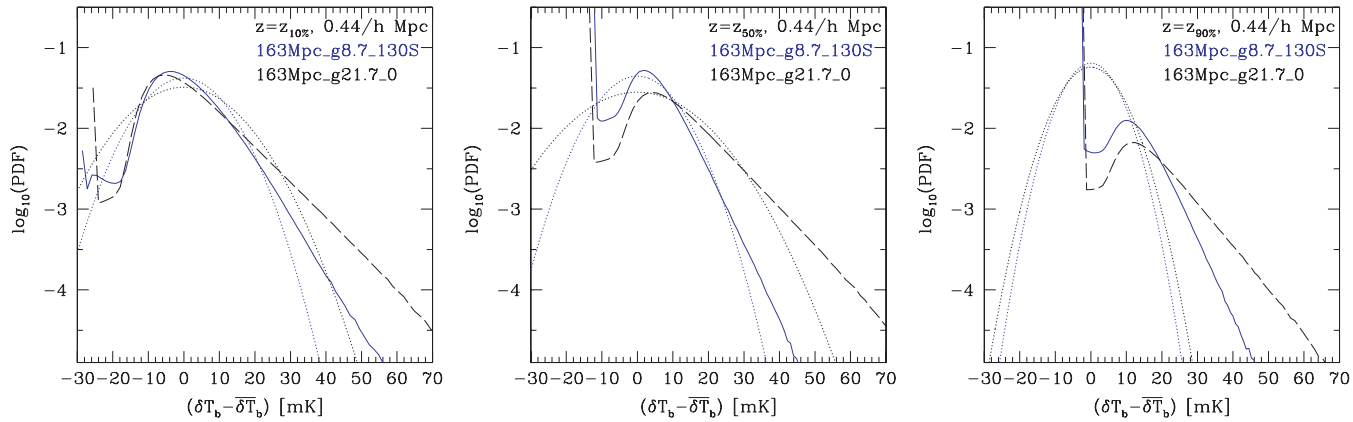

Figure 24. The effect of self-regulation on the PDF distribution of the 21-cm signal. Shown are the epochs at which the ionized fractions are $x_m = 0.1$ (left), $x_m = 0.5$ (middle) and $x_m = 0.9$ (right) for our fiducial self-regulated case, 163Mpc_g8.7_130S (L1; blue, solid) and the corresponding non-self-regulated case with the same overlap epoch, 163Mpc_g21.7_0 (L3; black, long-dashed). The PDFs are cell-by-cell (i.e. no smoothing apart from the numerical grid resolution). Also indicated are the Gaussian distributions with the same mean values and standard deviations (dotted lines, corresponding colours).

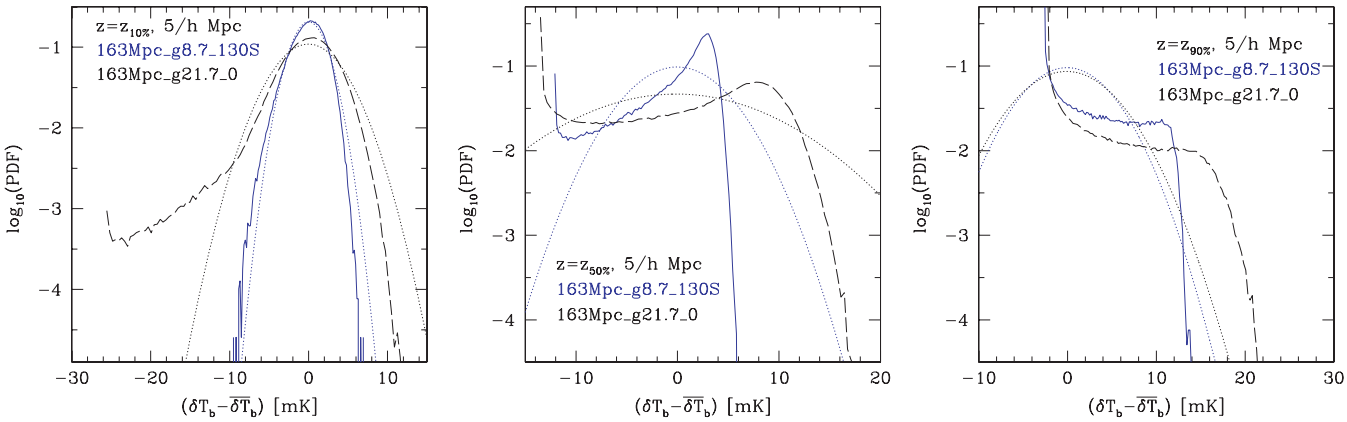

Figure 25. Same as in Fig. 24, but for boxcar smoothing of $5\,h^{-1}$ Mpc.

become much less wide compared to the unsmoothed ones, since the smoothing window averages the values, flattening the highest peaks and deepest valleys of the distributions.

Interestingly, at the middle and late stages of reionization ($x_m = 0.5$ and 0.9) the opposite happens, namely that the smoothed PDF distributions become less Gaussian for any δT_b value. The PDF distributions with and without self-regulation have similar shapes, but the presence of LMACHs makes the distribution much less wide. For $x_m = 0.5$, the very brightest peaks are fewer than a Gaussian would predict, but there are many more intermediate-brightness ($5\,\mathrm{mK} < \delta T_b - \delta\bar{T}_b < 12\,\mathrm{mK}$ with no self-regulation, $\delta T_b - \delta\bar{T}_b < 5\,\mathrm{mK}$ with) peaks. At the late stages of reionization ($x_m = 0.9$), both cases show many more bright peaks ($10\,\mathrm{mK} < \delta T_b - \delta\bar{T}_b$) than a Gaussian would predict, although the self-regulated case yields fewer very bright peaks ($15\,\mathrm{mK} < \delta T_b - \delta\bar{T}_b$) than either the corresponding Gaussian or the non-self-regulated case. Finally, regardless of the above differences in the PDFs, their equivalent widths are very similar for the two simulations.

The PDFs for our fiducial self-regulated high- and low-efficiency cases L1, L2, S1 and S2 are shown in Figs 26 and 27. Unlike the presence and self-regulation of LMACHs presented above, which influenced the PDFs significantly, neither the source efficiencies nor the box size has any dramatic effect on the resulting PDFs. The smaller boxes do not capture well the bright wing of the distribution because the highest density peaks are rare and the volume in these cases is too small to capture them. The effect of varying source efficiency manifests itself by yielding more bright peaks during the early stages of reionization and somewhat brighter peaks at its middle stages.

Finally, the results with varying source models are shown in Figs 28 and 29. Here for clarity we just show a representative subsample of our full simulation suite. Upon inspection, several general trends become clear. If only HMACHs are present (model S5) the distributions are noticeably wider, with a long non-Gaussian tail at high differential brightness temperatures ($\delta T_b - \delta\bar{T}_b > 30$–$40$ mK) than in the fiducial self-regulation case S1. Conversely, there are many fewer regions with low, but positive (i.e. still mostly neutral) differential brightness temperatures ($\delta T_b - \delta\bar{T}_b < 15$ mK). The reason for this is that the massive sources form only at the highest density peaks, leaving neutral many other density peaks which have not yet collapsed. The high-density, neutral gas in those peaks is reflected in the non-Gaussian tail at high differential brightness temperatures. Lastly, model S4 (low-efficiency sources and no suppression) yields intermediate PDF between the fiducial run and the HMACHs-only runs. The PDFs are therefore mostly dependent on which population of sources is active (HMACHs or LMACHs), but are not very sensitive to the details of the reionization history (models S5 and S8, not shown here, which have the same source population active, but with different efficiencies over time and thus different reionization history yield very similar distributions). These trends are independent of the smoothing employed, as can be seen in Fig. 29, although naturally the range of differential brightness temperatures is much reduced by the smoothing. The only new feature found in the smoothed data is the non-Gaussian tails for negative $\delta T_b - \delta\bar{T}_b$ at early times ($x_m = 0.1$). These are a result of the H II regions ($\delta T_b - \delta\bar{T}_b < 0$) in the non-self-regulated cases growing relatively large quickly. Consequently, even at these early times, their sizes become comparable to the smoothing window size, which results in the non-Gaussian tails. These were not present in the cell-wise PDF distribution, as the individual cells tend to be either fully ionized ($\delta T_b - \delta\bar{T}_b < -25$ mK) or mostly neutral.

The level of non-Gaussianity of the PDF distributions can, to a first order, be characterized by their skewness, which in turn can be used to distinguish and extract the reionization signals (Harker et al. 2009). In Fig. 30 we show the evolution of the skewness versus frequency for selected models. We show the skewness for the cell-wise PDF, as well as smoothed with a 3-arcmin Gaussian beam and 440-kHz bandwidth and smoothed with Gaussian beam corresponding to a 2.5-km maximum baseline and 440-kHz bandwidth (bottom panels). Both sets of beam and bandwidth smoothing are roughly as expected for the LOFAR array. We also plotted the rms of the correspondingly smoothed 21-cm differential brightness temperature fluctuations (top panels). We note that because of the different beamwidth and bandwidth smoothing employed here the rms values are slightly different from those shown in Figs 17, 19 and 20.

In all cases, regardless of the specific reionization scenario, the skewness evolution for the unsmoothed (1-cell) PDFs follows a similar pattern. It stays at a roughly constant, positive value throughout the evolution until it shoots up just before overlap, very similar to the behaviour observed in Harker et al. (2009). The beamwidth and bandwidth smoothing of the PDFs introduces a significant feature in the skewness, whereby it becomes negative during the intermediate stages of the evolution. Interestingly, this dip of the skewness to negative values closely corresponds to the rise and peak of the differential brightness temperature rms fluctuations, preceding it slightly in time. This feature is universal, observed for every reionization scenario and source model we consider here and suggests an interesting approach for a detection and/or independent

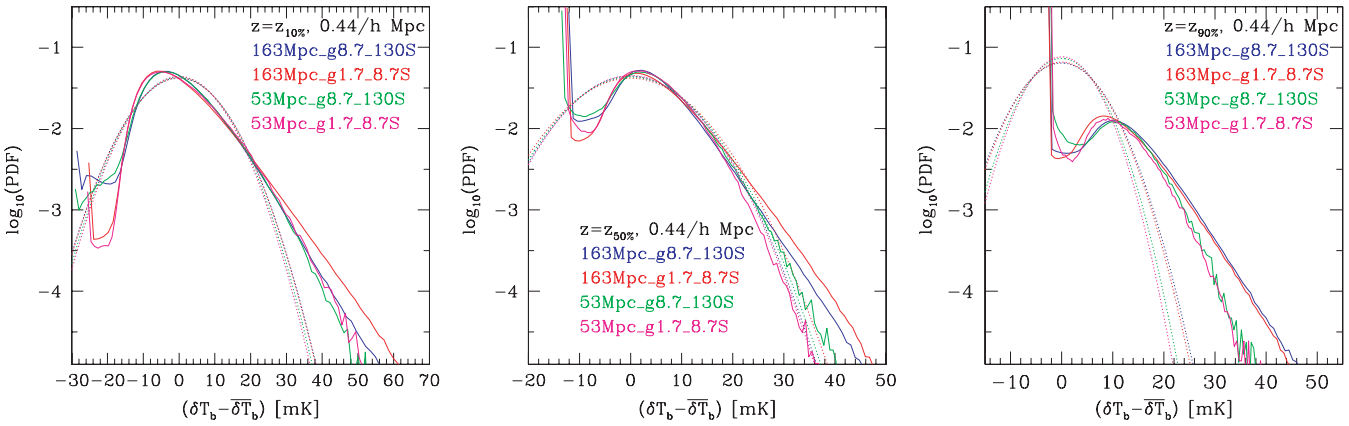

Figure 26. The effect of the source efficiencies (high versus low efficiency) and box size on the PDF distribution of the 21-cm signal. Shown are the epochs at which the ionized fractions are $x_m = 0.1$ (left), $x_m = 0.5$ (middle) and $x_m = 0.9$ (right) for our fiducial self-regulated cases, 163Mpc_g8.7_130S (L1; blue, solid), 163Mpc_g1.7_8.7S (L2; red, solid), 53Mpc_g8.7_130S (S1; green, solid) and 53Mpc_g1.7_8.7S (S2; magenta, solid). The PDFs are cell-by-cell (i.e. no smoothing apart from the numerical grid resolution). Also indicated are the Gaussian distributions with the same mean values and standard deviations (dotted lines, corresponding colours).

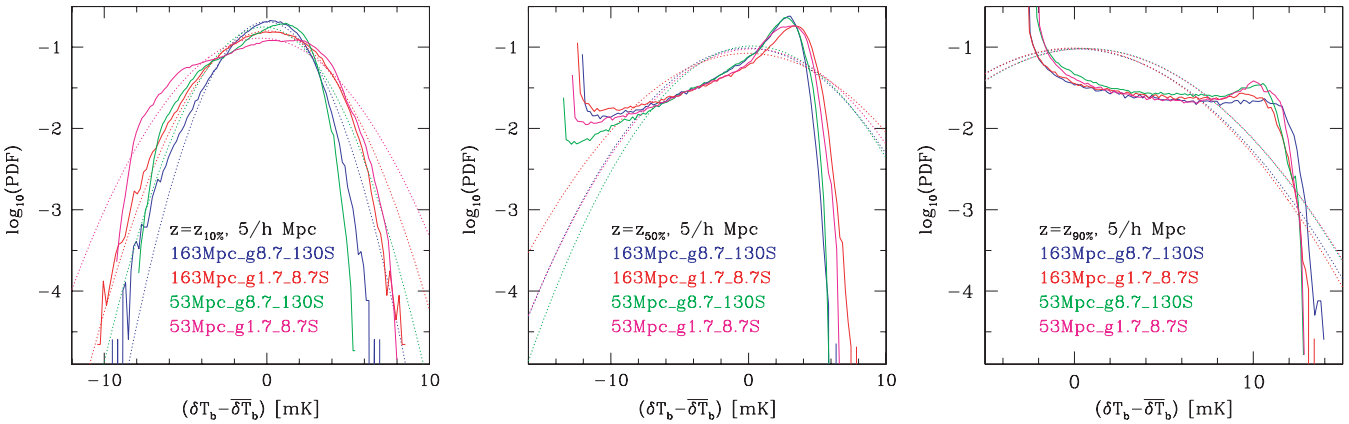

Figure 27. Same as in Fig. 26, but for boxcar smoothing of $5\,h^{-1}$ Mpc.

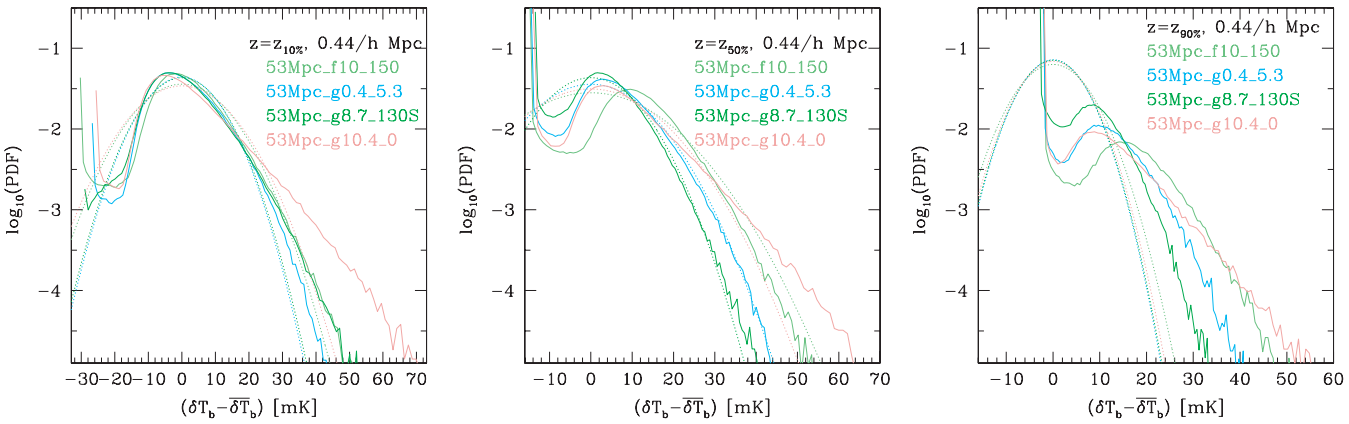

Figure 28. The effect of varying source models on the PDF distribution of the 21-cm signal. Shown are the epochs at which the ionized fractions are $x_m = 0.1$ (left), $x_m = 0.5$ (middle) and $x_m = 0.9$ (right) for our fiducial self-regulated case, 53Mpc_g8.7_130S (S1; green, solid), 53Mpc_g0.4_5.3 (S4; cyan, solid) and 53Mpc_g10.4_0 (S5; light red, solid). The PDFs are cell-by-cell (i.e. no smoothing apart from the numerical grid resolution). Also indicated are the Gaussian distributions with the same mean values and standard deviations (dotted lines, corresponding colours).

confirmation of the rise and peak of the 21-cm rms fluctuations during cosmic reionization.

On the other hand, the skewness of the 21-cm PDF distributions proves to be fairly insensitive to the source model or reionization scenario, resulting in only slight changes in the values. The skewness of the smoothed PDFs also proves largely independent of the box size and resolution and of the details of the interferometer beam assumptions (i.e. if it is fixed in angular size or evolves with

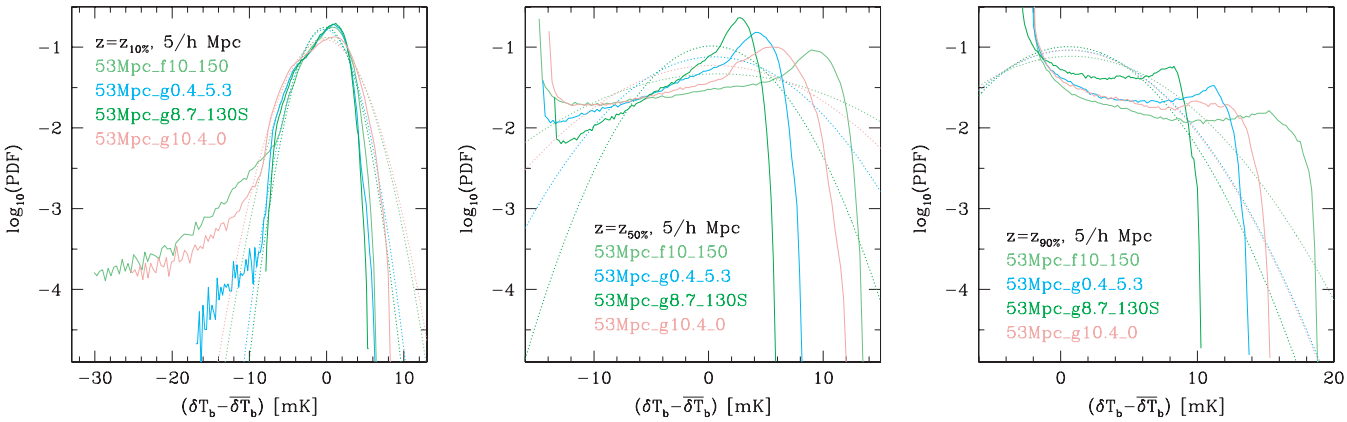

Figure 29. Same as in Fig. 28, but for boxcar smoothing of $5\,h^{-1}$ Mpc.

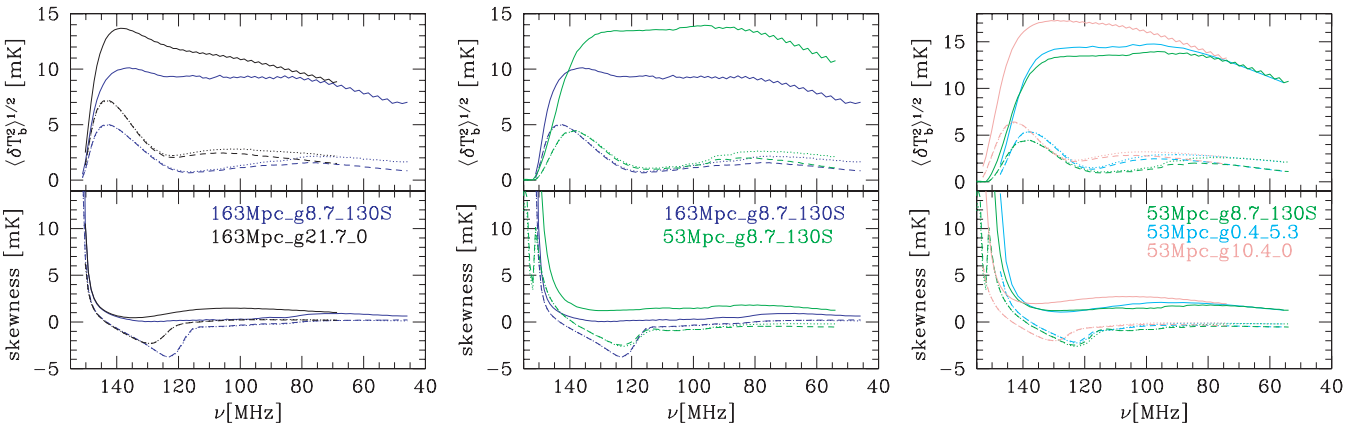

Figure 30. Top: the evolution of the rms of the 21-cm fluctuations; bottom: evolution of the skewness of the 21-cm PDFs for (left-hand panels) simulations 163Mpc_g8.7_130S (L1; blue) and 163Mpc_g21.7_0 (L3; black); (middle panels) simulations L1 (blue) and 53Mpc_g8.7_130S (S1; green); and (right-hand panels) simulations S1 (green), 53Mpc_g0.4_5.3 (S4; cyan) and 53Mpc_g10.4_0 (S5; light red). Shown are the results at full simulation resolution (solid lines), smoothed with a 3-arcmin Gaussian beam and 440 kHz bandwidth (dotted lines) and smoothed with a Gaussian beam corresponding to a 2.5-km maximum baseline and 440-kHz bandwidth (dashed lines).

frequency). This suggests that while the feature in the evolution of the skewness is a good indicator of the rise in patchiness, it most likely cannot be used for constraining the properties of the ionizing sources.

6 SUMMARY

In this work we have used a large set of cosmological structure formation and reionization simulations in an attempt to gain insight into what can be learned about the properties of the reionization sources based on their observational signatures. In particular, we are interested in determining what observations could be used to discriminate between certain source models, thereby restricting the available parameter space. Here we primarily focused on the redshifted 21-cm signatures, as these can in principle probe the full reionization history and offer a wide range of different probes, from the mean history, through fluctuation measures like rms evolution and power spectra, to PDFs and higher order statistics which can detect non-Gaussian features.

The observable features of the epoch of reionization derive from the gradual mean transition of the IGM from neutral to a highly ionized state, as well as from the patchiness of that transition. The mean transition depends largely on the overall number of ionizing photons emitted by the sources per unit time, with some correction due to recombinations which is position dependent due to density spatial variations. The patchiness, on the other hand, depends in a complicated way on the abundances, clustering and efficiencies of the ionizing sources.

Our structure formation simulations confirm previous results by Iliev et al. (2006a), Reed et al. (2007), Lukić et al. (2007) and Cohn & White (2008) that the high-redshift halo mass functions are inconsistent with either of the widely used Press–Schechter and Sheth–Tormen analytical fits. In particular, the abundance of rare haloes is strongly underestimated by PS, but overestimated by ST. We find that the non-linear halo bias is extremely high and very scale dependent. The linear bias regime is only reached at very large scales, $k \gtrsim 0.1\,h\,\mathrm{Mpc}^{-1}$. For the rarest haloes (3σ and above) the linear bias regime is never reached even within our largest $114/h = 163$ Mpc volume. Therefore, proper account for the non-linear bias of haloes is important and any calculations assuming linear bias are underestimating the halo clustering significantly.

The Jeans mass filtering of low-mass haloes (LMACHs) in ionized regions and the related self-regulation of reionization results in a significantly more extended reionization history and a higher

integrated electron-scattering optical depth (by $\Delta\tau_{es} \sim 0.01$) compared to the high-mass source (HMACHs)-only scenario with the same overlap redshift, albeit both optical depths are still within the current constraints from the *WMAP* observations. Even more significant are the changes in the reionization geometry, resulting in corresponding differences in the reionization observables. Compared to the HMACH-only scenario, all self-regulated cases yield 21-cm rms fluctuations which are lower at all scales and PDF distributions which are somewhat more Gaussian, although significant non-Gaussianity remains.

In all our simulations reionization occurs inside-out, with the high-density regions being reionized on average earlier than the mean and low-density ones. This inside-out nature of reionization results in the mass-weighted IGM photoionization rates being considerably larger, by factor of a few, than volume-weighted ones. This should always be taken into account as it can easily skew observational measurements depending on the mean density of the regions being probed.

The skewness of the 21-cm PDF distribution smoothed over LOFAR-like beam and bandwidth shows a clear feature correlated with the rise of the rms due to patchiness. This feature does not exist in the unsmoothed data, indicating that it is related to the non-Gaussianity of the large-scale patchy distribution of 21-cm emission. The feature exists for any reionization scenario and ionizing source properties and thereby provides a different approach for detecting the rise of large-scale patchiness and an independent check on other detections.

The peak position of the 21-cm rms fluctuations depends significantly on the beam- and bandwidth smoothing size as well as on the reionization scenario. As a consequence, it does not always occur at a 50 per cent ionization fraction as is sometimes claimed, but instead can happen for ionized fractions as low as 30 per cent and as high as 70 per cent.

The ionizing source efficiencies and their correlation properties introduce clear signatures in the reionization observables. As a direct consequence of that, one cannot model low-mass unresolved sources by simply assigning their emissivity to the resolved higher mass sources. The latter have an abundance which has steeper time evolution and different clustering properties from lower mass sources which provide the bulk of the ionizing photons (see Appendix A).

When self-regulation is present there are only minor differences between the 21-cm observational signatures resulting from high- and low-efficiency ionizing sources, apart from an overall shift of the reionization history. The corresponding PDF distributions are also very similar, which suggests that the source efficiencies in such models can only be constrained by the overall timing of the mean reionization history.

Scenarios where LMACHs are completely absent, e.g. if they are somehow rendered sterile, are relatively easily distinguishable from the ones where they are present (even if they happen to be strongly suppressed). On the other hand, our results suggest that numerous low-efficiency sources (case S4) can mimic the effects of suppression (S1). Such scenarios therefore will be difficult to distinguish solely based on power spectra and similar measurements. However, they might still be discriminated through 21-cm PDFs as the no-suppression case creates many more high-brightness 21-cm peaks. Similarly, an HMACHs-only scenario (S5) gives quite similar results to the high-efficiency, no LMACH suppression case (S3) at the same stages of reionization (albeit these cases do reach overlap at different times for the parameters we have chosen), although they do differ in terms of power at large scales and in number of bright 21-cm peaks.

In this work (and a companion paper, Friedrich et al. 2011) we for the first time considered the effects of simulation volume and radiative transfer numerical grid resolution on the reionization properties and observables. The simulation volume has only a modest effect on the results as long as the typical size of the ionized patches is smaller than the volume. However, the fluctuations at large scales (above approximately a fifth of the box size) are severely affected. This is especially important at late times when the ionized patches grow very large. Therefore, at least $\sim 100/h = 143$ Mpc per side simulation boxes are required to model correctly the fluctuations during the late stages of reionization.

The effects of the finite numerical grid resolution of the radiative transfer simulation are more complex and vary significantly depending on the quantity being considered. The mean reionization history (and mean 21-cm emission) is insensitive to moderate variation of the resolution. The small-scale density fluctuations can potentially increase the local mean recombination rate and thus affect the reionization history and properties. However, resolving them requires either much higher spatial resolution or subgrid modelling, both of which are beyond the scope of this work. The 21-cm fluctuations are largely insensitive to the numerical resolution as long as the beam/bandwidth of the radio array is resolved and the ionized regions are larger than the array smoothing scale. This is typically the case throughout most of the reionization history, but not at the earliest times, when the ionized regions are still small and thus structures are averaged out by insufficient numerical or observational resolution. Similarly, the 21-cm brightness PDFs are insensitive to resolution if smoothed at scales larger than a few numerical cells. In contrast, other reionization features like the 21-cm PDF non-Gaussianity (skewness) and the topological properties of the ionized regions (Friedrich et al. 2011) are sensitive to smoothing.

The results presented in this work should not be considered to be an ultimate, realistic prediction of the reionization signals. While the assumptions about source efficiencies and suppression we made for our fiducial cases are reasonable based on our current knowledge and likely bracket the realistic range, the uncertainties are still substantial. As more observational data become available over time, these can be used to restrict the parameter space further and help us refine our theoretical models, which in turn will provide a valuable tool for interpreting the meaning of the observational results in terms of early structure formation, HMACH/LMACH source efficiencies, LMACH suppression mechanisms, etc. In this framework our current study, which evaluates in a controlled way the effects of a set of widely different assumptions about the sources of ionizing radiation, is a very useful step towards a more complete understanding of early galaxy formation and feedback.

ACKNOWLEDGMENTS

ITI was supported by The Southeast Physics Network (SEPNet) and the Science and Technology Facilities Council grants ST/F002858/1 and ST/I000976/1. This study was supported in part by Swiss National Science Foundation grant 200021-116696/1, Swedish Research Council grant 2009-4088, NSF grants AST-0708176 and AST-1009799, NASA grants NNX07AH09G, NNG04G177G and NNX11AE09G, and Chandra grant SAO TM8-9009X. The authors acknowledge the TeraGrid and the Texas Advanced Computing Center (TACC) at The University of Texas at Austin (URL: http://www.tacc.utexas.edu), as well as the Swedish

National Infrastructure for Computing (SNIC) resources at HPC2N (Umeå, Sweden) for providing HPC and visualization resources that have contributed to the research results reported within this paper and Partnership for Advanced Computing in Europe (PRACE) grant 2010PA0442 to ITI. KA is supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (MEST; 2009-0068141, 2009-0076868) and by KICOS through K20702020016-07E0200-01610 funded by MOST.

REFERENCES

Ade P. A. R. et al. (Planck Collaboration), 2011, A&A, 536, 1
Ahn K., Shapiro P. R., Iliev I. T., Mellema G., Pen U., 2009, ApJ, 695, 1430
Aubert D., Teyssier R., 2010, ApJ, 724, 244
Baek S., di Matteo P., Semelin B., Combes F., Revaz Y., 2009, A&A, 495, 389
Bowman J. D., Rogers A. E. E., 2010, Nat, 468, 796
Bowman J. D., Morales M. F., Hewitt J. N., 2007, ApJ, 661, 1
Ciardi B., Ferrara A., White S. D. M., 2003, MNRAS, 344, L7
Cohn J. D., White M., 2008, MNRAS, 385, 2025
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Cucchiara A. et al., 2011, ApJ, 736, 7
Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12
Doré O., Holder G., Alvarez M. A., Iliev I. T., Mellema G., Pen U.-L., Shapiro P. R., 2007, Phys. Rev. D, 76, 043002
Efstathiou G., 1992, MNRAS, 256, 43P
Fernandez E. R., Komatsu E., Iliev I. T., Shapiro P. R., 2010, ApJ, 710, 1089
Field G. B., 1959, ApJ, 129, 536
Friedrich M. M., Mellema G., Alvarez M. A., Shapiro P. R., Iliev I. T., 2011, MNRAS, 413, 1353
Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, ApJ, 613, 1
Furlanetto S. R., McQuinn M., Hernquist L., 2006a, MNRAS, 365, 115
Furlanetto S. R., Oh S. P., Briggs F. H., 2006b, Phys. Rep., 433, 181
Gluscevic V., Barkana R., 2010, MNRAS, 408, 2373
Gnedin N. Y., 2000, ApJ, 535, 530
Gnedin N. Y., Jaffe A. H., 2001, ApJ, 551, 3
Haiman Z., Abel T., Rees M. J., 2000, ApJ, 534, 11
Harker G. J. A. et al., 2009, MNRAS, 393, 1449
Harker G. et al., 2010, MNRAS, 405, 2492
Ichikawa K., Barkana R., Iliev I. T., Mellema G., Shapiro P. R., 2010, MNRAS, 406, 2521
Iliev I. T., Shapiro P. R., Raga A. C., 2005, MNRAS, 361, 405
Iliev I. T., Mellema G., Pen U.-L., Merz H., Shapiro P. R., Alvarez M. A., 2006a, MNRAS, 369, 1625
Iliev I. T. et al., 2006b, MNRAS, 371, 1057
Iliev I. T., Mellema G., Shapiro P. R., Pen U.-L., 2007a, MNRAS, 376, 534
Iliev I. T., Pen U.-L., Bond J. R., Mellema G., Shapiro P. R., 2007b, ApJ, 660, 933
Iliev I. T., Mellema G., Pen U., Bond J. R., Shapiro P. R., 2008a, MNRAS, 384, 863
Iliev I. T., Shapiro P. R., McDonald P., Mellema G., Pen U.-L., 2008b, MNRAS, 391, 63
Iliev I. T., Shapiro P. R., Mellema G., Merz H., Pen U.-L., 2008c, preprint (arXiv:0806.2887)
Iliev I. T. et al., 2009, MNRAS, 400, 1283
Iliev I. T., Ahn K., Koda J., Shapiro P. R., Pen U.-L., 2010, preprint (arXiv:1005.2502)
Kashikawa N. et al., 2011, ApJ, 734, 119
Komatsu E. et al., 2011, ApJS, 192, 18
Krug H. B. et al., 2012, ApJ, 745, 122
Larson D. et al., 2011, ApJS, 192, 16
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Liu A., Tegmark M., 2012, MNRAS, 419, 3491
Lonsdale C. J. et al., 2009, Proc. IEEE, 97, 1497
Lukić Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, ApJ, 671, 1160
McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, ApJ, 653, 815
McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, MNRAS, 377, 1043
Mao Y., Shapiro P. R., Mellema G., Iliev I. T., Koda J., Ahn K., 2012, MNRAS, in press
Mellema G., Iliev I. T., Alvarez M. A., Shapiro P. R., 2006a, New Astron., 11, 374
Mellema G., Iliev I. T., Pen U.-L., Shapiro P. R., 2006b, MNRAS, 372, 679
Merz H., Pen U.-L., Trac H., 2005, New Astron., 10, 393
Mesinger A., Dijkstra M., 2008, MNRAS, 390, 1071
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Mortlock D. J. et al., 2011, Nat, 474, 616
Nakamoto T., Umemura M., Susa H., 2001, MNRAS, 321, 593
Navarro J. F., Steinmetz M., 1997, ApJ, 478, 13
Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920
Ouchi M. et al., 2010, ApJ, 723, 869
Paciga G. et al., 2011, MNRAS, 413, 1174
Parsons A. R. et al., 2010, AJ, 139, 1468
Press W. H., Schechter P., 1974, ApJ, 187, 425
Quinn T., Katz N., Efstathiou G., 1996, MNRAS, 278, L49
Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2007, MNRAS, 374, 2
Ricotti M., Gnedin N. Y., Shull J. M., 2002, ApJ, 575, 33
Rijkhorst E.-J., Plewa T., Dubey A., Mellema G., 2006, A&A, 452, 907
Salvaterra R., Ciardi B., Ferrara A., Baccigalupi C., 2005, MNRAS, 360, 1063
Shapiro P. R., Giroux M. L., Babul A., 1994, ApJ, 427, 25
Shapiro P. R., Iliev I. T., Raga A. C., 2004, MNRAS, 348, 753
Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, A&A, 345, 380
Sheth R. K., Tormen G., 2002, MNRAS, 329, 61
Shin M.-S., Trac H., Cen R., 2008, ApJ, 681, 756
Sokasian A., Abel T., Hernquist L., Springel V., 2003, MNRAS, 344, 607
Susa H., Umemura M., 2004, ApJ, 610, L5
Thomas R. M. et al., 2009, MNRAS, 393, 32
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Trac H., Cen R., 2007, ApJ, 671, 1
White M., 2001, A&A, 367, 27
Zahn O., Mesinger A., McQuinn M., Trac H., Cen R., Hernquist L. E., 2011, MNRAS, 414, 727

APPENDIX A: REIONIZATION BY RARE, MASSIVE, VARIABLE-LUMINOSITY SOURCES

Simulations 53Mpc_uvS_1e9 and 53Mpc_uvS_1e10 (S8 and S9 in Table 2) are investigating the effects of keeping the global, volume-averaged emissivity of ionizing photons per unit time fixed at each redshift, while raising the minimum source mass, assumed to be $10^9\,M_\odot$ for S8 and $10^{10}\,M_\odot$ for S9. The overall number of photons emitted at each timestep are set to be exactly equal, at all times, to the one yielded by our fiducial case, S1. The resulting effective efficiencies g_γ are shown in Fig. A1. Simulation S8 has the same HMACH population as our fiducial simulation, but no active LMACHs at all. Therefore, the effective efficiencies start very high, at several thousand photons per atom, as the relatively few HMACHs at early time have to 'compensate' for the more numerous LMACHs present in the fiducial case, as well as for all photons emitted before $z \sim 19$ in our fiducial simulation, during which time there are no active sources larger than $10^9\,M_\odot$. However, as the number of HMACHs rises exponentially, $g_{\gamma,\mathrm{eff}}$ drops precipitously, to less than 40 by $z = 12.6$ and less than 20 by $z = 11$. Towards overlap $g_{\gamma,\mathrm{eff}}$ settles on $\sim$8.7, the value adopted in our fiducial case, as by then all LMACHs are suppressed and the HMACHs are identical to those

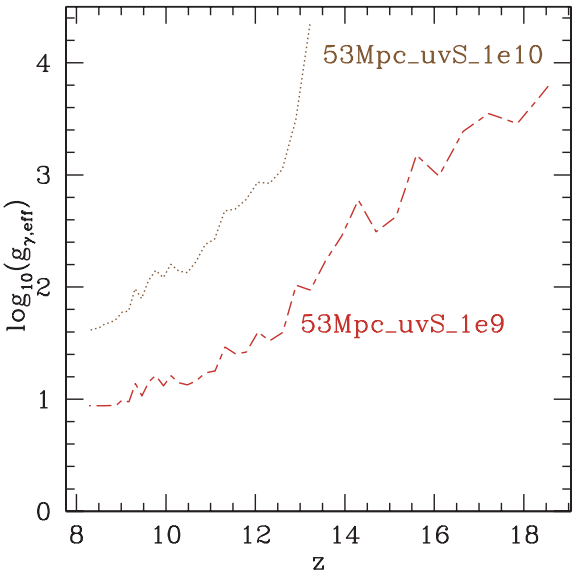

Figure A1. Effective efficiency factors g_γ versus redshift for simulations 53Mpc_uvS_1e9 (S8) and 53Mpc_uvS_1e10 (S9), defined so as to ensure the same total number of ionizing photons emitted per atom as in our fiducial simulation 163Mpc_g8.7_130S (S1).

in the fiducial case. Note that although in this case the sources belong to the same haloes as in S5 and overlap is reached at a similar redshift, this S8 case is different in assuming the same step-by-step total emissivity as in our fiducial case, L1, which naturally makes them variable in time, unlike case L5 discussed before, which had a fixed photon emissivity per unit halo mass.

Case S9 is still more extreme, since only quite massive haloes, with masses above $10^{10}\,M_{\odot,}$ are allowed to be active sources (i.e. even the smaller HMACHs are assumed to be suppressed). The first such massive haloes form in our simulation only at $z = 13.2$ and they remain relatively rare ($\sim 3\sigma$) even at overlap ($z = 8.2$). As a consequence, their effective efficiency is very high at all times, starting at over 20 000 and reaching $\sim$40 at overlap. In order to avoid hyperluminous sources during the first timestep, we distributed the photons that were emitted at $z > 13.2$ in the fiducial case over the first several timesteps of run S9. Clearly, both S8 and S9 scenarios are not very realistic physically, given this vast range of change in the source efficiencies.

Since in cases S8 and S9 we imposed the same global integrated ionizing photon emissivities per timestep as in our fiducial case S1, the averaged global reionization histories of those two cases closely follow that of the fiducial simulation once the first haloes above the respective minimum cut-off form in our volume. The only remaining difference is that at early times ($z > 11$) the ionized fraction in case S9 is a little lower than in the other two cases, as a consequence of our imposition of a more gradual initial release of photons in this case, in order to avoid hyperluminous sources, as explained above. However, unlike S1, both simulations S8 and S9 yield $x_{\rm m}/x_{\rm v} \approx 1$, since in those latter cases the ionized patches produced by the few, luminous sources present are far less correlated with the underlying density field (Fig. A2). This is a consequence of the I-fronts quickly escaping into the nearby voids, which compensates for the exponential rise of the number of ionizing sources forming at the high-density peaks (although we note that even in this case reionization remains inside-out, as the ionized regions are still overdense on average). This results in H II region distributions which are clearly distinct from all other cases considered. As the cut-off mass increases, there are exponentially fewer ionizing sources, which consequently are much more efficient (cf. Fig. A1). Hence, those hyperluminous sources produce correspondingly large H II regions, which are less correlated with the underlying density field and are more spherical than in the other cases, as they are produced by few, but highly clustered sources.

This distinct H II region geometry of cases S8 and S9 also yields very characteristic 21-cm signatures (Fig. A3). The massive, rare, highly efficient sources quickly produce very extended H II regions and thus high rms fluctuations at large scales and a very broad peak, with an almost constant value ($\langle \delta T_{\rm b}^2 \rangle^{1/2} = 10.39$–$10.94$) for a wide range of mean ionized fractions by mass, $x_{\rm m} = 0.21$–0.58. There is also no initial dip of the rms fluctuations, which normally occurs when the highest density peaks are ionized, but the H II regions are still much smaller than the smoothing beam size. In models S8 and S9, the ionized patches grow so fast that their typical sizes are of the order of or larger than the beam at all times. Such a scenario therefore yields a signal which is both stronger and quite different from the others. The results for the lower minimum source mass cut-off case with the same reionization history, S8, show similar properties to S9, namely a broad and relatively high rms peak and no initial dip. However, the peak value in this case, at $\sim$8 mK, is noticeably lower than that for model S9 ($\sim$11 mK) and is more similar to the typical values for the majority of cases ($\sim$5–8 mK).

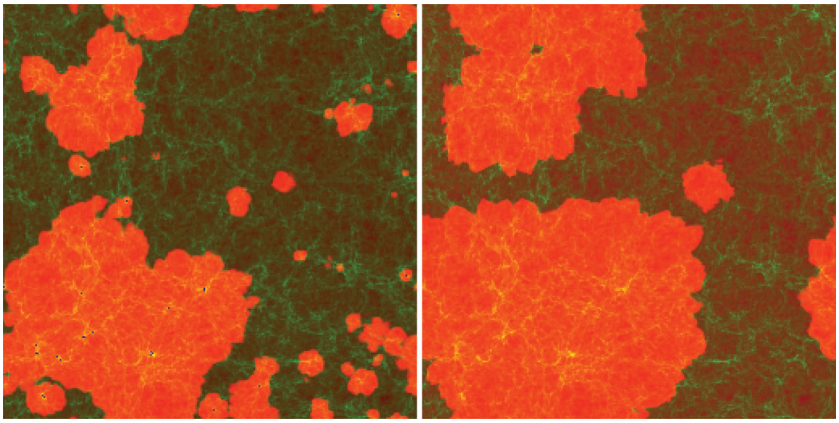

Figure A2. Spatial slices of the ionized and neutral gas density from our radiative transfer simulations with box size 53 Mpc at a box-averaged ionized fraction by mass $x_{\rm m} \sim 0.50$. Shown are the density field (green) overlaid with the ionized fraction (red/orange/yellow) and the cells containing sources (dark/blue). Shown are cases 53Mpc_uvS_1e9 (S8) and 53Mpc_uvS_1e10 (S9).

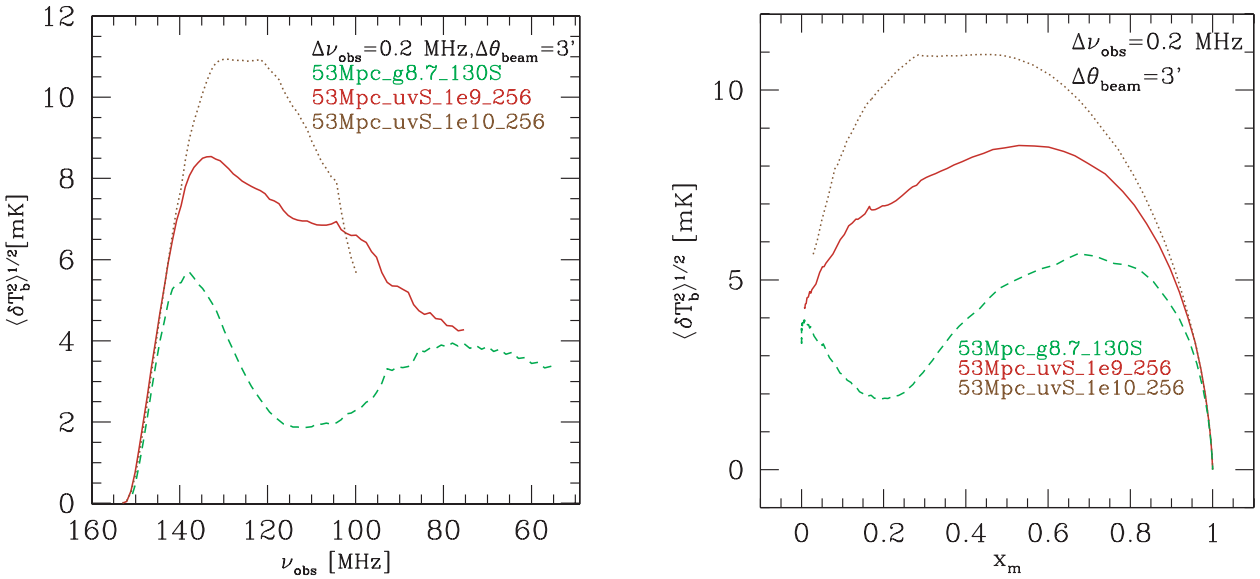

Figure A3. The evolution of the rms fluctuations of the 21-cm background, for beamsize 3 arcmin and bandwidth 0.2 MHz versus frequency (left) and versus mean mass-weighted ionized fraction (right). Shown are simulations 53Mpc_g8.7_130S (S1; green, short dashed), 53Mpc_uvS_1e9 (S8; dark red, long dash-short dash) and 53Mpc_uvS_1e10 (S9; brown, dotted).

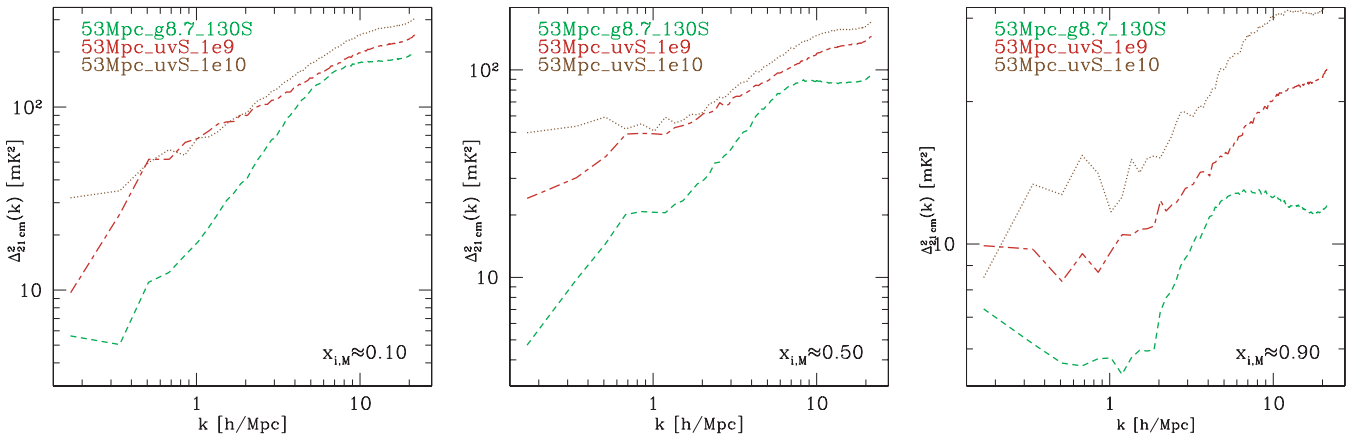

Figure A4. The 21-cm differential brightness temperature fluctuation power spectra for varying source models. Shown are the epochs at which the ionized fractions are (left) $x_m = 0.1$, (middle) $x_m = 0.5$ and (right) $x_m = 0.9$. All cases are labelled by colour and line type, as follows: 53Mpc_g8.7_130S (S1; green, short-dashed), 53Mpc_uvS_1e9 (S8; dark red, long dash-short dash) and 53Mpc_uvS_1e10 (S9; brown, dotted).

The corresponding PDF distributions (not shown) are noticeably wider, with a long non-Gaussian tail at high differential brightness temperatures ($\delta T_b - \delta\bar{T}_b > 30$–$40$ mK) than in the fiducial self-regulation case S1.

The corresponding 21-cm power spectra (Fig. A4) for cases S8 and S9 also show a significantly higher signal on all scales compared to our fiducial case S1. During the early stages of reionization, the power at large scales ($k \lesssim 0.4\,h\,\mathrm{Mpc}^{-1}$) for S9 is almost an order of magnitude higher than that for S1. During the late stages, the difference decreases considerably, but still remains $\sim$2 on large ($k < 2\,h\,\mathrm{Mpc}^{-1}$), as well as very small ($k > 8\,h\,\mathrm{Mpc}^{-1}$) scales. As could be expected, the case with lower minimum mass, S8, is intermediate between S1 and S9, but much closer to S9 throughout the evolution.

Our scenario S9 is similar to the high minimum source mass case, S4, considered in McQuinn et al. (2007). These authors set the minimum source mass to $4 \times 10^{10}\,\mathrm{M}_\odot$, somewhat higher than that in S9. Their ionizing photon production is similarly set to reproduce, step-by-step, the one of their fiducial case. Their results are qualitatively similar to that we find. The rare, efficient and strongly clustered sources yielded 21-cm power spectra which were higher and flatter than in their fiducial case, with the difference decreasing over time (cf. fig. 17 in McQuinn et al. 2007). However, some quantitative differences remain, due to the somewhat different approach we have taken, as well as some numerical and resolution differences. Apart from the higher source mass cut-off adopted by McQuinn et al. (2007), which results in a stronger source bias, other important differences include lower resolution of their *N*-body and radiative transfer simulations, and the lack of Jeans mass filtering. Unlike our high-resolution simulations, which resolve all atomically cooling haloes ($M > 10^8\,\mathrm{M}_\odot$), the *N*-body structure formation simulations used by McQuinn et al. (2007) resolved only haloes with mass above $10^9\,\mathrm{M}_\odot$, with LMACHs included in some cases by subgrid modelling. More importantly, their fiducial case (whose photon production per timestep was the basis for their high-mass cut-off case S4) yielded late overlap and included no Jeans mass

filtering (several of their other simulations included it, but not this one). Therefore, their photon production per unit source mass was necessarily very low, making their fiducial case more similar to our low-efficiency case S4 than to our fiducial simulation S1. Finally, we take account of peculiar velocity when calculating more precise 21-cm power spectra (including redshift-space distortions, Mao et al. 2012). Despite these differences, our results agree reasonably well on a qualitative level.

We note that models S8 and S9 are rather unrealistic, as they assume unphysically high and time-variable luminosities, as well as the suppression of all sources with mass below $10^{10}\,M_{\odot}$ (or, less aggressively, $10^{9}\,M_{\odot}$ for case S8), for which no clear mechanism exists. We have included these models here primarily in order to demonstrate, under controlled circumstances, the effect of higher source-mass cut-off on the 21-cm observables. Such a higher source cut-off mass occurs numerically in simulations with large volumes and limited dynamic range (e.g. Baek et al. 2009; Thomas et al. 2009) and, therefore, it is important to evaluate the level of reliability of such models. Our results show that including only the high-mass sources can result in overestimating the 21-cm rms fluctuations by up to a factor of 2, while $P(k)$ at small k where the first generation of observations will probe, could be overestimated by as much as an order of magnitude at the 50 per cent ionized epoch. It can also yield quite a different evolution, even for the same box size, numerical resolution and the same integrated photon emissivity over time. One should therefore be aware of these potential pitfalls and adjust their modelling accordingly. A better simulation approach would be to add the unresolved LMACHs by subgrid modelling.

APPENDIX B: DEPENDENCE ON THE JEANS SUPPRESSION THRESHOLD

Here we consider variations of our source suppression threshold, with the goal of establishing the robustness and validity of our fiducial model, S1, where we use $x_{\rm threshold} = 0.1$. We ran two additional models, S6 and S7, in which we raised this ionization threshold for LMACHs suppression to $x_{\rm threshold} = 0.9$ and 0.5, respectively. In our fiducial case, the suppression criterion for partially ionized cells is more aggressive than in these new cases. The reasonable value to be adopted for this suppression threshold is still quite uncertain at present; thus, it is important to check the sensitivity of our results to variations in its value. In fact, it is most likely that a sharp on-off condition like this is an oversimplification and in reality the suppression boundary is gradual, with full suppression of the smallest galaxies, partial one for intermediate-mass galaxies, up to no suppression at all for sufficiently massive galaxies. However, given the current uncertainties, the range of possible suppression models is very large and it is difficult to fully explore numerically. Instead, we have chosen to consider three very different cases covering the full range of the threshold value, in order to evaluate the effect of these uncertainties on the reionization history and observables. We note, however, that we consider our original source suppression criterion to be well motivated, for the following reasons. Although for numerical reasons our suppression criterion is defined in terms of the ionized fraction, physically it is related to the temperature state of the IGM, for which the ionization state is used as a proxy. When a given region is photoionized, its temperature rises to $\sim 10^4$ K, with the exact value dependent on the intensity and spectrum of the ionizing radiation, ranging from $\sim$20 000 K for Pop. II stellar spectra and QSOs to $\sim$30 000–40 000 K for Pop. III (Shapiro et al. 2004). This rises the gas pressure and thus the Jeans mass to $10^{9}\,M_{\odot}$ or more. In order for the LMACHs to be able to re-form in a previously ionized region its temperature should decrease to well below 10^4 K. However, in the mostly metal-free gas during these early epochs, there is no efficient radiative coolant available and therefore the main cooling mechanisms are the local adiabatic expansion and Compton scattering of CMB photons. Since both of these processes are relatively slow and inefficient, we expect that our fiducial more aggressive ionized fraction-based source suppression criterion is more physically realistic than the milder suppression of the new cases. However, given the significant uncertainties of the Jeans filtering process, which can only be properly modelled by hydrodynamical simulations with detailed and realistic microphysics, we consider all of these very different suppression criteria and study their consequences below.

The reionization histories and cumulative numbers of ionizing photons emitted derived for the three suppression criteria are shown in Fig. B1 (left). Clearly, only a very weak suppression ($x_{\rm threshold} = 0.9$, case S6) yields any significant differences. Compared to our fiducial case S1, many fewer LMACHs are suppressed in S6, and of these a significant fraction are allowed to become active again shortly after suppression (since in the absence of radiation recombinations quickly bring the neutral fraction back up above 10 per cent). The evolution of the number of photons produced in simulation S6 is up to $\sim$2 higher in the middle stages of reionization, while the corresponding number for S7 is essentially the same as in S1 throughout the evolution.

The Jeans mass filtering nonetheless still has a significant effect, keeping the ionized fraction well below the corresponding one for the no-suppression case S3 (cf. Fig. 9, right-hand panels). Eventually, the fully ionized fraction of the volume becomes sufficiently large to suppress almost all LMACHs even with this mild suppression criterion and the reionization process slows down until sufficient number of HMACHs form and are able to finish this process and reach overlap. In contrast, the intermediate case, S7 ($x_{\rm threshold} = 0.5$), shows only modest differences from the fiducial model S1, manifesting themselves mostly in bringing reionization forward by $\Delta z \sim 0.4$, compared to $\Delta z \sim 1.2$–3.2, and very different shape of the reionization history for case S6. Similarly, the integrated electron-scattering optical depth for the mild suppression case S6 is $\tau_{\rm es} = 0.111$, much higher than in the fiducial case ($\tau_{\rm es} = 0.080$), while for the intermediate suppression case the increase is much more modest, at $\tau_{\rm es} = 0.089$. The cumulative number of photons per atom at overlap, $\sim$2, is very similar in all three cases.

The variations in the geometry of reionization (Fig. B2) are mostly found in the small-scale structures. There are significantly fewer such structures in the weak suppression case S6. The merged H II regions are typically slightly larger, as well as rounder and with smoother boundaries compared to the fiducial simulation S1. Once again, the intermediate case S7 is very similar to S1, with only minor differences in small-scale features. These visual impressions are further confirmed by comparing the 21-cm power spectra (Fig. B3). The weak suppression case S6 has significantly more power at intermediate and large scales ($k \lesssim 5$) during all stages of reionization, more so at late times, but less power on small scales than our fiducial case S1. On the other hand, the intermediate model S7 matches S1 fairly closely, except for having less power on very small scales.

The 21-cm mean differential brightness temperature (Fig. B4, left) for the weak suppression case S6 shows an initial steep decline around $\nu \sim 100$ MHz, followed by a sudden change of slope at $\nu \sim 125$ MHz and a very slow decrease thereafter. This behaviour is quite different from models S1 and S7 (which again follow almost identical evolution), as well as from all other models discussed earlier. The only model with a similarly sharp decrease of the mean

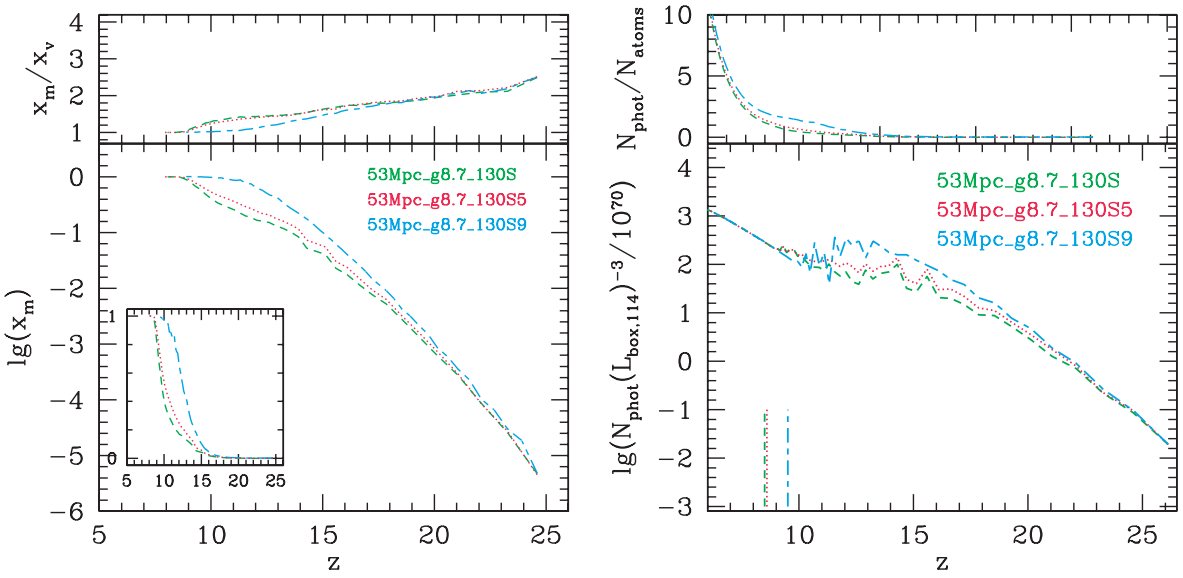

Figure B1. Left, bottom panel: Mass-weighted reionization histories for cases S1, S6 and S7, each with different Jeans mass filtering threshold. Left, top panel: ratio of the corresponding mean mass-weighted and volume ionized fractions, x_m/x_v. Right, bottom panel: the number of ionizing photons emitted by all active sources (thick lines) in the computational volume per timestep; and (right, top panel) cumulative number of photons per total gas atom released into the IGM. Vertical lines mark the overlap redshift in each case. All curves on both left and right are labelled by colour and line type, as follows: 53Mpc_g8.7_130S (S1; green, short-dashed), 53Mpc_g8.7_130S9 (S6; light blue, long dash-short dash) and 53Mpc_g8.7_130S5 (S7; red, dotted).

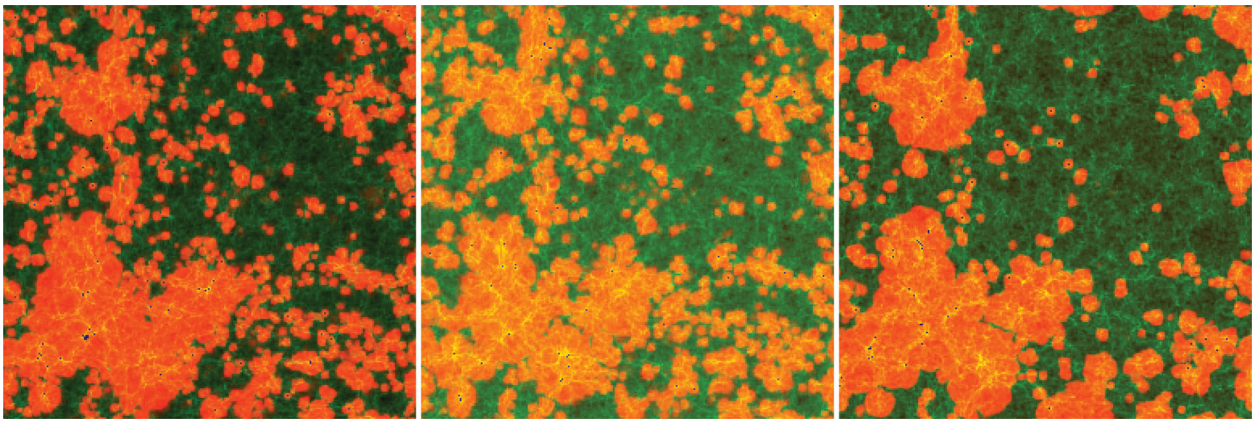

Figure B2. Spatial slices of the ionized and neutral gas density from our radiative transfer simulations with box size $37\,h^{-1}$ Mpc, all at box-averaged ionized fraction by mass $x_m \sim 0.50$. Shown are the density field (green) overlaid with the ionized fraction (red/orange/yellow) and the cells containing sources (dark/blue). Shown are cases 53Mpc_g8.7_130S (S1), 53Mpc_g8.7_130S5 (S7) and 53Mpc_g8.7_130S9 (S6).

brightness temperature is the high-efficiency, no-suppression case S3, which however does not have the same long slow evolution tail at late times due to lack of suppression.

Finally, the differential brightness temperature rms fluctuations (Fig. B4, right) for the weak suppression case, S7, peak much earlier, at $\nu \approx 110$ MHz (but still significantly later than the no-suppression case S3, which underlines the importance of even a very weak LMACHs suppression) reaching $\langle \delta T_b^2 \rangle^{1/2} \approx 8$ mK. Uniquely, this model exhibits a very long tail of slow decline of the rms fluctuations beyond the peak. This is related to its very different (and, as we argued earlier, possibly less physically realistic) suppression model. Case S6 also exhibits significant fluctuations in the differential brightness temperature fluctuations, once again indicating that this suppression model might be less physically realistic than our standard suppression model. The intermediate suppression model S6 follows the same evolution as the fiducial case S1, but shifted to slightly earlier time.

In summary, all results prove fairly insensitive to the precise value of the Jeans suppression threshold assumed, as long as it is not at the very weak suppression limit. Both $x_{threshold} = 0.1$ and 0.5 yield essentially the same evolution, apart from a slight offset in time. On the other hand, a very high suppression threshold ($x_{threshold} = 0.9$, i.e. weak suppression) results in a very different (and somewhat unstable) evolution with several characteristic observational signatures. We have, however, argued above that such a weak suppression is likely less realistic physically. Our suppression model therefore proves quite robust to a threshold variation within the plausible range.

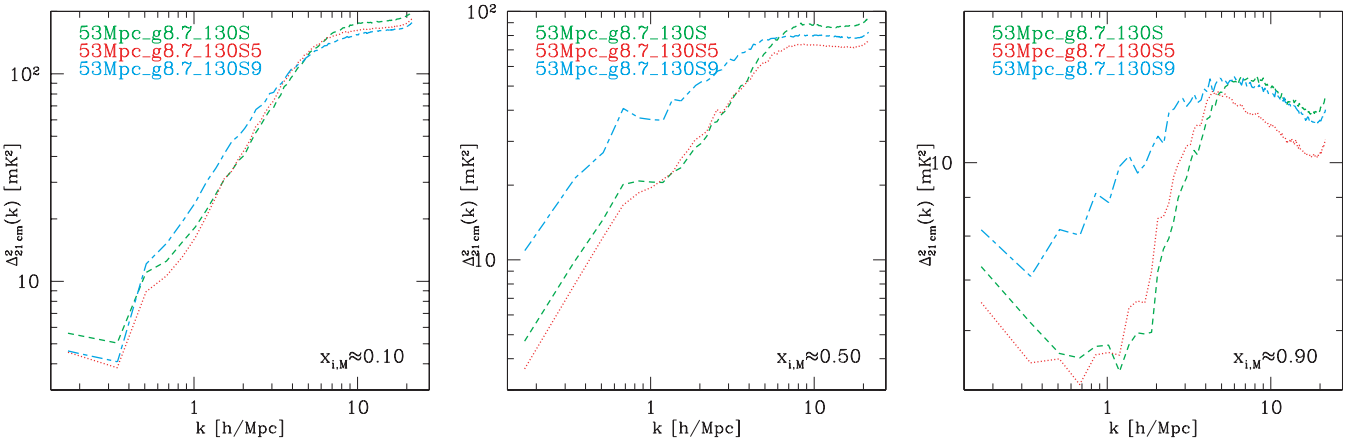

Figure B3. The 21-cm differential brightness temperature fluctuation power spectra for varying source models. Shown are the epochs at which the ionized fractions are (left) $x_m = 0.1$, (middle) $x_m = 0.5$ and (right) $x_m = 0.9$. All cases are labelled by colour and line type, as follows: 53Mpc_g8.7_130S (S1; green, short-dashed), 53Mpc_g8.7_130S9 (S6; light blue, long dash–short dash) and 53Mpc_g8.7_130S5 (S7; red, dotted).

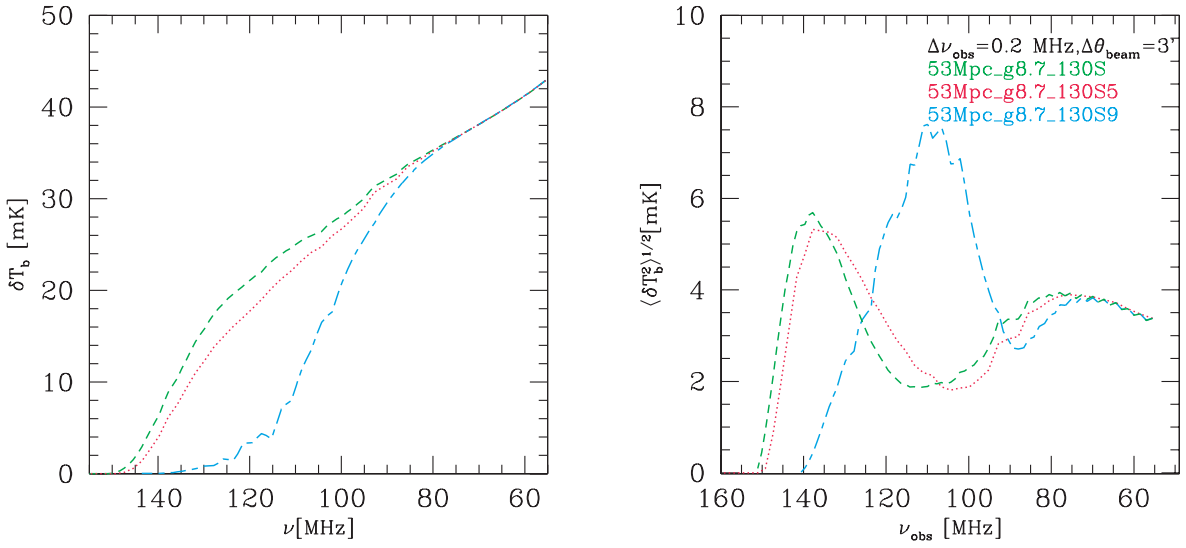

Figure B4. The evolution of the mean 21-cm background (left) and its rms fluctuations for Gaussian beamsize 3 arcmin and bandwidth 0.2 MHz (right) versus observed 21-cm frequency. Shown are simulations 53Mpc_g8.7_130S (S1; green, short-dashed), 53Mpc_g8.7_130S9 (S6; light blue, long dash-short dash) and 53Mpc_g8.7_130S5 (S7; red, dotted).

This paper has been typeset from a TEX/LATEX file prepared by the author.